ESAIM: Probability and Statistics Will be set by the publisher
URL: http://www.emath.fr/ps/

ON A SZEGÖ TYPE LIMIT THEOREM, THE HÖLDER-YOUNG-BRASCAMP-LIEB INEQUALITY, AND THE ASYMPTOTIC THEORY OF INTEGRALS AND QUADRATIC FORMS OF STATIONARY FIELDS *

Florin Avram[1], Nikolai Leonenko[2] and Ludmila Sakhno[3]

Abstract. Many statistical applications require establishing central limit theorems for sums/integrals $S_T(h) = \int_{t\in I_T} h(X_t)dt$ or for quadratic forms $Q_T(h) = \int_{t,s\in I_T} \hat{b}(t-s)h(X_t, X_s)dsdt$, where X_t is a stationary process. A particularly important case is that of Appell polynomials $h(X_t) = P_m(X_t)$, $h(X_t, X_s) = P_{m,n}(X_t, X_s)$, since the "Appell expansion rank" determines typically the type of central limit theorem satisfied by the functionals $S_T(h), Q_T(h)$. We review and extend here to multidimensional indices, along lines conjectured in [16], a functional analysis approach to this problem proposed by Avram and Brown (1989), based on the method of cumulants and on integrability assumptions in the spectral domain; several applications are presented as well.

Résumé. Nous considérons ici des théorémes de limite central pour des sommes/integrales $S_T(h) = \int_{t\in I_T} h(X_t)dt$, et pour des formes quadratiques $Q_T(h) = \int_{t,s\in I_T} \hat{b}(t-s)h(X_t, X_s)dsdt$, où X_t est un processus stationnaire. Un cas particulièrement important est celui des polynômes d'Appell $h(X_t) = P_m(X_t)$, $h(X_t, X_s) = P_{m,n}(X_t, X_s)$. Pour ce problème, nous généralisons ici au cas des indices multidimensionnels une approche proposée par Avram et Brown (1989), basée sur la méthode des cumulants et sur des hypothèses d'integrabilité dans le domaine spectral. Plusieurs applications illustrent la versatilité de l'approche.

1991 Mathematics Subject Classification. 60F05, 62M10, 60G15, 62M15, 60G10, 60G60.

The dates will be set by the publisher.

Contents

Keywords and phrases: quadratic forms, Appell polynomials, Hölder-Young inequality, Szegö type limit theorem, asymptotic normality, minimum contrast estimation.

* *Partly supported by the grant of London Mathematical Society 2632, the EPSRC grant RCMT 119 and by the Welsh Institute of Mathematics and Computational Sciences.*

[1] Dept. de Math., Universite de Pau, France, E-mail: Florin.Avram@univ-Pau.fr
[2] Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, UK,
Email: LeonenkoN@Cardiff.ac.uk
[3] Dept. of Probability Theory and Mathematical Statistics, Kyiv National Taras Shevchenko University, Ukraine, and Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, UK, Email:lms@univ.kiev.ua

Introduction

Model. For a unified treatment of the discrete and the continuous (multi-dimensional) setups, we assume that $X_t, t \in I$ is a real stationary random field, where I denotes a measurable group endowed with its Haar measure, and $\int_{t\in I} \nu(dt)$ is integral with respect to the Haar measure. Depending on the continuous/discrete setup, I will be either $\mathbb{R}^d$ with Lebesgue measure, or $\mathbb{Z}^d$ with the counting measure. Discrete sums might be written either as integrals (in the statement of theorems), or in traditional sum notation (in the expository part). For continuous case we assume that $X_t, t \in \mathbb{R}^d$ is a measurable random field.

We will assume throughout the existence of all order cumulants $c_k(t_1, t_2, ..., t_k)$ for our stationary in the strict sense random field X_t, which are supposed to be representable as Fourier transforms of "cumulant spectral densities" $f_k(\lambda_1, ..., \lambda_{k-1}) \in L_1$, i.e:

$$c_k(t_1, t_2, ..., t_k) = c_k(t_1 - t_k, .., t_{k-1} - t_k, 0) =$$

$$= \int_{\lambda_1,...,\lambda_{k-1}\in S} e^{i\sum_{j=1}^{k-1} \lambda_j(t_j - t_k)} f_k(\lambda_1, ..., \lambda_{k-1})\, \mu(d\lambda_1)...\mu(d\lambda_{k-1}).$$

Throughout, S will denote the "spectral" space of discrete/continuous processes and $\mu(d\lambda)$ will denote the associated Haar measure, i.e. Lebesgue measure normalized to unity on $[-\pi, \pi]^d$ and Lebesgue measure on $\mathbb{R}^d$, respectively. The functions $f_k(\lambda_1, ..., \lambda_{k-1})$ are symmetric and generally complex valued for real field X_t.

The random field X_t is observed on a sequence I_T of increasing dilations of a finite convex domain I_1, i.e.

$$I_T = T\, I_1, \quad T \to \infty.$$

Correspondingly, we will consider linear and bilinear forms S_T and Q_T, where summation/integration is performed over domains I_T.

In the discrete-time case, the cases $I_T = [1,T]^d$, $T \in \mathbb{Z}_+$ (in keeping with tradition) or $I_T = [-T/2, T/2]^d$, $T \in 2\mathbb{Z}_+$ will be assumed. In the continuous case, we focus on rectangles $I_T = \{t \in \mathbb{R}^d : -T/2 \le t_i \le T/2, i = 1, ..., d\}$.

Later on we will see that the choice of a specific set I_1 leads (when working in the spectral domain) to using an associated Dirichlet type kernel

$$\Delta_T(\lambda) = \int_{t \in I_T} e^{it\lambda} \nu(dt) \tag{0.1}$$

and a multivariate Fejér kernel (0.2). Explicit well-known formulas are available when I_1 is a rectangle or a ball, both for the discrete and continuous case. We will work in the paper with rectangular domains, however, some extensions are possible for balls by replacing corresponding estimates (see Appendix B).

Motivation. Via the work of Szego, Schur, Wiener and Kolmogorov, the study of stationary one-dimensional discrete time series, defined by choosing $I = \mathbb{Z}$, has been well developed and became tightly interwoven with several branches of mathematics, bringing forth jewels like for example the Wiener-Kolmogorov formula identifying the variance of the prediction error with respect to the past with the integral of the log of the spectral density. The extension to the continuous time case $I = \mathbb{R}$, provided by Krein, exemplifies the subtle challenges brought by modifying the nature of the index set.

The convenience of time series comes largely from the FARIMA family of parametric models, defined as solutions of equations

$$\phi(B)(1-B)^d X_t = \theta(B)\xi_t$$

where B is the operator of backward translation in time, $\phi(B), \theta(B)$ are polynomials, d is a real number and ξ_t is white noise [51], [45], [50], [23], [73]. Using the FARIMA family of models, one attempts, via an extension of the Box-Jenkins methodology, to estimate the parameter d and the coefficients of the polynomials $\phi(B), \theta(B)$ such that the residuals ξ_t display white noise behavior (and hence may be safely discarded for prediction purposes).

An extension of this approach to continuous time and to multi-parameter processes (spatial statistics) has been long time missing. Only recently, the FICARMA ([28], [6]) and the Riesz-Bessel families of processes (which essentially replace the operator B by the Laplacian - see Appendix A), have allowed a similar approach for processes with continuous and multidimensional indices (at least in the rotationally invariant case).

These examples illustrate the point that extension of ARIMA-type models to continuous and multidimensional indices is an important challenge.

Central limit theorems for quadratic forms. Asymptotic statistical theory, and in particular, estimation of the parameters of FARIMA and Riesz-Bessel processes, requires often establishing central limit theorems concerning

$$\text{sums} \qquad S_T = S_T(h) = \sum_{i=1}^{T} h(X_i)$$

$$\text{and bilinear forms} \qquad Q_T = Q_T(h) = \sum_{i=1}^{T} \sum_{j=1}^{T} \hat{b}(i-j)\ h(X_i, X_j).$$

of functions of stationary sequences X_i (and their generalizations to continuous and multidimensional indicies).

Notes: 1) The kernel of the quadratic form will be denoted by $\hat{b}(t)$, so that we may reserve $b(\lambda)$ for its Fourier transform.

2) The random field X_t, $t \in I_T$, will be allowed to have short-range or long-range dependence (that is, summable or non-summable correlations).

A particularly important case is that of Appell polynomials[1] $h(X_t) = P_m(X_t)$, $h(X_t, X_s) = P_{m,n}(X_t, X_s)$ associated to the distribution of X_t, which are the building block of the so called "chaos/Fock expansions". Two main cases were distinguished, depending on whether the limit is Gaussian or not (the latter case being referred to as a non-central limit theorem).

This line of research, initiated by Dobrushin and Major [32] and Taqqu [71] in the Gaussian case (see also Ivanov and Leonenko (1989) for Gaussian continuous case), by Giraitis and Surgailis [40], [44], and by Avram and Taqqu [11] in the linear case, continues to be of interest today [38], [42], [43], [60], [63], [67] to name only a few papers.

Our interest here is in obtaining extensions to continuous and multi-parameter processes of the central limit theorem for sums and quadratic form, obtained in the case of discrete time series by Breuer and Major [27], by the method of moments.

Some analytic tools. A key unifying role in our story is played by the multivariate Fejér kernels:

$$\Phi_T^*(u_1, ..., u_{n-1}) = \frac{1}{(2\pi\mu(I_1))^{(n-1)d}\, T^d} \Delta_T(-\sum_1^{n-1} u_e) \prod_{e=1}^{n-1} \Delta_T(u_e), \tag{0.2}$$

and their kernel property: the fact that **when $T \to \infty$, the multivariate Fejér kernel Φ_T^* convergence weakly to a δ measure**:

Lemma 0.1. ***The kernel property:*** *For any continuous bounded function $C(u_1, .., u_{n-1})$, it holds that:*

$$\lim_{T\to\infty} \int_{S^{n-1}} C(u_1, .., u_{n-1}) \Phi_T^*(u_1, ..., u_{n-1}) \prod_{i=1}^{n-1} \mu(du_i) = C(0, ..., 0)$$

Proof: For the discrete onedimensional case we refer for example to [12], and for the continuous multidimensional case, with I_1 a rectangle, to Bentkus [21], [22] or [8], Proposition 1.[2].

Developing some limit theory for multivariate Fejér kernels was the key point in several papers [12], [14], [13], [15] which generalized the Breuer and Major central limit theorem [27]. The papers above introduced a new mathematical object to be called "**Fejér graph kernels**" – see (2.4) in section 2, which captures the common structure of several cumulant computations. Replacing the cyclic graph encountered in the specific case of quadratic forms in Gaussian random variables by an arbitrary graph, these papers reduce the central limit theorem for a large class of problems involving Appell polynomials in Gaussian or moving average summands to an application of three analytical tools:

(1) Identifying the graphs involved by applying the well-known **diagram formula** for computing moments/cumulants of Wick products – see section E.

[1]For the definition of Appell polynomials see, for instance, Avram and Taqqu (1987), or Giraitis and Surgailis (1986).

[2]This convergence of measures may also be derived as a consequence of the Hölder-Young-Brascamp-Lieb inequality (see Theorem C.1), using estimates of the form

$$||\Delta_T||_{s_v^{-1}} \leq k(s_v) T^{d(1-s_v)}$$

with optimally chosen s_v, $v = 1, \cdots, V$.

(2) Applying a generalization of a **Grenander-Szegö theorem on the trace of products of Toeplitz matrices** to the Fejér graph integrals – see section 2, to obtain the asymptotic variance. This theorem is valid under some general integrability assumptions furnished by the Hölder-Young-Brascamp-Lieb inequality.

(3) The resolution of certain combinatorial graph optimization problems, specific to each application, which clarify the geometric structure of the polytope of valid integrability exponents on the functions involved (spectral density, kernel of the quadratic form, etc).

Here, we observe that a similar approach works in the multidimensional and continuous indices case. More precisely, the only changes are a) the normalizations, which change from T to T^d, and b) the condition for the validity of the Hölder-Young-Brascamp-Lieb inequality (see Appendix C) in the continuous case. Therefore, the previously obtained central limit theorems continue to hold in the multidimensional case, including continuous indices, after simply adjusting the normalizations and integrability conditions.

Statistical applications. The two cases most easy to study are that of Gaussian and linear processes. In the applications Section 4 we will work assuming that X_t is a linear process (see (4.1) or (4.3) below). This assumption has the advantage of implying a product representation of the cumulant spectral densities – see for example Theorem 2.1 of [6]. Namely, for the cumulants we get the explicit formula

$$c_k(t_1, ..., t_k) = d_k \int_{s \in I} \prod_{j=1}^{k} \hat{a}(t_j - s) \, \nu(ds),$$

and in the spectral domain, we get

$$f_k(\lambda_1, ..., \lambda_{k-1}) = d_k \; a(-\sum_{i=1}^{k-1} \lambda_i) \; \prod_{i=1}^{k-1} a(\lambda_i) = d_k \; \prod_{i=1}^{k} a(\lambda_i) \delta(\sum_{j=1}^{k} \lambda_j) \tag{0.3}$$

(the meaning of parameters d_k and a function $a(\lambda)$ is clarified in section 4.1).

For $k = 2$, we will denote the spectral density by $f(\lambda) = f_2(\lambda) = d_2 a(\lambda) a(-\lambda)$.

Definition 0.2. Let

$$\mathbf{L}_p(d\mu) = \begin{cases} L_p(d\mu) & \text{if } p < \infty, \\ C & \text{if } p = \infty. \end{cases} \tag{0.4}$$

denote the closure of the functions in L_p which are **continuous, bounded and of bounded support**, under the L_p norm.

Note: In the torus case, this space intervenes in the proof of theorem 3.1, which is established first for complex exponentials, and extended then to the Banach space of functions which may be approximated arbitrarily close in $L_p(d\mu)$ sense by linear combinations of complex exponentials, endowed with the L_p norm.

Considering bilinear forms Q_T we will work under integrability assumption:

Assumption A:

$$f \in \mathbf{L}_{p_1}(d\mu), b \in \mathbf{L}_{p_2}(d\mu), \quad 1 \le p_i \le \infty, i = 1, 2$$

Note: 1) While a general stationary model is parameterized by a sequence of functions $f_k(\lambda_1, ..., \lambda_{k-1})$, $k \in \mathbb{N}$, the linear model (0.3) is considerably simpler, being parametrized by a single function $a(\lambda)$.

2) We expect all our results may be formulated directly in terms of characteristics of the field X_t, which suggests that the moving average assumption is probably unnecessary; indeed, more general results which make direct assumptions that functions $f_k(\lambda_1, ..., \lambda_{k-1})$ belong to some special L_p-type spaces, have been obtained in certain cases – see, for example, [15].

Contents. We present a warmup example involving quadratic Gaussian forms in Section 1. The results here are closely connected to those of the paper by Ginovian and Sahakyan (2007). We define the concept of Fejér graph and matroid integrals in Section 2. We will consider here only the first case (i.e. graphic matroids associated to the incidence matrix of a graph).

Some limit theory (of Grenander-Szegö type) for Fejér graph integrals is reviewed and extended to the continuous case in Section 3. Various estimates concerning kernels are collected in Appendix B, and a particular case of the Hölder-Young-Brascamp-Lieb inequality required here is presented in Appendix C.

In Section 4 we introduce the linear model (which extends the Gaussian model) and develop several applications. Note here the existence of a different approach, due to Peligrad and Utev (2006), who established the central limit theorem for linear processes with discrete time and dependent innovations including martingale and mixingale type assumptions (see also the references therein for this line of investigation).

To make the paper self-contained we supply in Appendices the material we refer to in the main part of the paper.

1. An example: the central limit theorem for Gaussian bilinear forms

We present first our method in the simplest case of symmetric bilinear forms $Q_T = Q_T^{(1,1)} = Q_T(P_{1,1})$ in stationary Gaussian fields X_t, with covariances $r(t-s)$, $t, s \in I$ and spectral density $f(\lambda)$ (note that $S_T^{(1)} = \int_{t \in I_T} X_t \nu(dt)$ is "too simple" for our purpose, since it is already Gaussian and and its k-th order cumulants $\chi_k(S_T^{(1)}) = 0$, $\forall k \neq 2$). The presentation follows [14] for the discrete case and [38] for the continuous case, except that we clarify the point that the previous results hold in any dimension d.

To obtain the central limit theorem for $T^{-d/2} Q_T^{(1,1)}$ by the method of cumulants it is enough to show that:

$$\lim_{T\to\infty} \chi_2 \left(\frac{Q_T^{(1,1)}}{T^{d/2}} \right) \text{ is finite, and } \quad \lim_{T\to\infty} \chi_k \left(\frac{Q_T^{(1,1)}}{T^{d/2}} \right) = 0, \ \forall k \geq 3.$$

A direct computation based on multilinearity yields the cumulants of $Q_T^{(1,1)}$:

$$\chi_k = \chi(Q_T, ..., Q_T) = 2^{k-1}(k-1)! \operatorname{Tr}[(T_T(b) T_T(f))^k]. \tag{1.1}$$

Here, Tr denotes the trace and

$$T_T(b) = (\hat{b}(t-s), \ t, s \in I_T), \qquad T_T(f) = (r(t-s), \ t, s \in I_T)$$

denote Toeplitz matrices (with multidimensional indices) of dimension $T^d \times T^d$ in the discrete case and truncated Toeplitz-type operators in the continuous case [3].

[3]Recall that in continuous case, the truncated Toeplitz-type operator generated by a function $\hat{f} \in L_\infty$, is defined for $u \in L_2$ as follows

$$T_T(f)u(t) = \int_{I_T} \hat{f}(t-s) u(s) \nu(ds).$$

While the cumulants in (1.1) may be expressed using powers of two Toeplitz matrices, it turns out more convenient in fact to consider more general products with all terms potentially different (taking advantage thus of multi-linearity).

Suppose therefore given a set $f_e(\lambda) : S \to \mathbb{R},\ e = 1, \ldots, n$ of "symbols" associated to the set of Toeplitz operators, where $(S, d\mu)$ denotes either $\mathbb{R}^d$ with Lebesgue measure, or the torus $[-\pi, \pi]^d$ with normalized Lebesgue measure. Assume the symbols satisfy integrability conditions

$$f_e \in \mathbf{L}_{p_e}(S, d\mu),\ 1 \le p_e \le \infty. \tag{1.2}$$

Let $\hat{f}_e(k), k \in I$, be the Fourier transform of $f_e(\lambda)$:

$$\hat{f}_e(k) = \int_S e^{ik\lambda} f_e(\lambda)\mu(d\lambda), \quad k \in I,$$

where $I = \mathbb{Z}^d$ in the torus case and $I = \mathbb{R}^d$ in the case $S = \mathbb{R}^d$, respectively. In this last case, we would also need to assume that $f_e \in L_1(\mathbb{R}^d, d\mu)$, for the Fourier transform to be well defined. Consider the extension of our cumulants:

$$\begin{aligned} \tilde{J}_T &= \mathrm{Tr}[\prod_i^n T_T(f_i)] \\ &= \int_{j_1,\ldots,j_n \in I_T} \hat{f}_1(j_2 - j_1)\hat{f}_2(j_3 - j_2)\ldots\hat{f}_n(j_1 - j_n) \prod_{v=1}^n \nu(dj_v). \end{aligned} \tag{1.3}$$

Replacing the sequences $\hat{f}_e(t)$ by their Fourier representations $\hat{f}_e(t) = \int_S f_e(\lambda)e^{it\lambda}d\lambda$ in (1.3) yields the following alternative spectral integral representation for traces of products of Toeplitz matrices or truncated Toeplitz operators[4]:

$$J_T = \int_{\lambda_1,\ldots,\lambda_n \in S} f_1(\lambda_1)f_2(\lambda_2)\ldots f_n(\lambda_n) \prod_{e=1}^n \Delta_T(\lambda_{e+1} - \lambda_e) \prod_{e=1}^n \mu(d\lambda_e) = \tag{1.4}$$
$$\int_{u_1,\ldots,u_{n-1} \in S} \Big(\int_{\lambda \in S} f_1(\lambda)f_2(\lambda + u_1)\ldots f_n(\lambda + \sum_1^{n-1} u_e)d\lambda \Big) \Phi_T(u_1, \ldots, u_{n-1}) \prod_{e=1}^{n-1} \mu(du_e),$$

where the index $n + 1$ is defined to be equal 1, and where

$$\Phi_T = \Delta_T(-\sum_1^{n-1} u_e) \prod_{e=1}^{n-1} \Delta_T(u_e),$$

which, after normalization with the factor $\frac{1}{(2\pi)^{(n-1)d}T^d}$, yields the "multivariate Fejér kernel"

$$\Phi_T^*(u_1, \ldots, u_{n-1}) = \frac{1}{(2\pi)^{(n-1)d}\, T^d} \Phi_T(u_1, \ldots, u_{n-1}).$$

---

[4]Of course, the two expressions $\tilde{J}_T, J_T$ are equal if $f_e \in L_1, e = 1, \ldots, n$. Note however that the "spectral representation" (1.4) is well defined even without the last condition.

Note that the inner integral

$$C(u) = C_{(f_1,...,f_n)}(u_1, ..., u_{n-1}) := \int_{\lambda \in S} f_1(\lambda) f_2(\lambda + u_1)...f_n(\lambda + \sum_1^{n-1} u_e) d\lambda, \tag{1.5}$$

to be called a graph convolution, is well defined precisely under the classical Hölder conditions, when the integrability indicies in (1.2) satisfy:

$$\begin{cases} \sum_e p_e^{-1} & \leq 1 \text{ when } S = \mathbb{Z}^d \\ \sum_e p_e^{-1} & = 1 \text{ when } S = \mathbb{R}^d, \end{cases} \tag{1.6}$$

The resulting integral

$$J_T = \int_{u_1,...,u_{n-1} \in S} C(u_1, ..., u_n) \Phi_T(u_1, ..., u_{n-1}) \prod_{e=1}^{n-1} \mu(du_e)$$

is our first example of a "Fejér graph integral" to be introduced in general in the next section. These are integrals involving products of Dirichlet kernels Δ_T and functions, applied to linear combinations, which are related to the vertex-edge incidence structure of a certain directed graph (in the occurrence, the cyclic graph on the vertices $\{1, ..., n\}$).

The second expression in the RHS of (1.4) reveals the asymptotic behavior of Fejér graph integrals, since, as noted, when $T \to \infty$, the multivariate Fejér kernel Φ_T^* **convergences weakly to a** δ **measure**.

The kernel lemma 0.1 will imply that

$$T^{-d} J_T \to (2\pi)^{(n-1)d} C(0, ..., 0) = \int_{\lambda \in S} f_1(\lambda) f_2(\lambda)...f_n(\lambda) \mu(d\lambda)$$

provided that we check that the function $C(u_1, ..., u_{n-1})$ is bounded and continuous. This is indeed true, as stated in the next result:

Lemma 1.1. *The "graph convolution" function $C(u_1, ..., u_{n-1})$ defined in (1.5) is bounded and continuous if (1.2) holds with integrability indices satisfying the condition (1.6).*

Proof: Note that the function $C : S^{n-1} \to \mathbb{R}$ is a composition

$$C(u_1, ..., u_{n-1}) = \mathcal{C}_{(T_1(u_1,...,u_{n-1}),...,T_n(u_1,...,u_{n-1}))}$$

of the functional

$$\mathcal{C}_{(f_1,...,f_n)} : \prod_{e=1}^{n} \mathbf{L}_{p_e} \to \mathbb{R}.$$

defined by [5]

$$\mathcal{C}_{(f_1,...,f_n)} := \int_{\lambda \in S} \prod_{e=1}^{n} f_e(\lambda) \mu(d\lambda)$$

[5]Note again this is well defined precisely under the classical Hölder conditions.

with the continuous functionals

$$T_e(u_1, ..., u_{n-1}) : S^{n-1} \to \mathbf{L}_{p_e}$$

defined by $T_e(u_1, ..., u_{n-1}) = f_e(\cdot + \sum_{v=1}^{e-1} u_v), e = 1, ..., n$.

Indeed, the continuity of the functionals T_e is clear when f_e is a function which is continuous, bounded and of bounded support, and this continues to be true for functions $f_e \in \mathbf{L}_{p_e}$, since these can be approximated in the L_{p_e} sense. In conclusion, under the Hölder assumptions, the continuity of the functional $C(u_1, ..., u_{n-1})$ will follow automatically from that of $\mathcal{C}_{(f_1,...,f_n)}$.

Finally, under the "Hölder conditions" (1.6), the continuity as well as boundedness of the function $\mathcal{C}_{(f_1,...,f_E)}$ follow from its multi-linearity and from the Hölder inequality:

$$|\mathcal{C}_{(f_1,...,f_n)}| \le \prod_{e=1}^{n} \|f_e\|_{p_e} \ .$$

Theorem 1.2. *Let $f_e \in \mathbf{L}_{p_e}, e = 1, ..., n$ where $1 \le p_e \le \infty, e = 1, ..., n$, and let $J_T, \tilde{J}_T$ be defined by (1.4), (1.3) respectively. Then, it follows that:*

(1) *If the integrability indices satisfy the Hölder conditions*

$$\begin{cases} \sum_e p_e^{-1} & \le 1 \text{ when } S = \mathbb{Z}^d \\ \sum_e p_e^{-1} & = 1 \text{ when } S = \mathbb{R}^d, \end{cases}$$

then

$$\lim_{T\to\infty} T^{-d} J_T = \int_{\lambda \in S} \prod_{e=1}^{n} f_e(\lambda)\mu(d\lambda) \tag{1.7}$$

(2) *If $\alpha := \sum_e p_e^{-1} > 1$, then it holds that:*

$$J_T = o(T^{\alpha d}).$$

(3) *If in the continuous case it holds in addition that if $f_e \in \mathbf{L}_{p_e} \cap L_1, e = 1, ..., n$, then Fourier coefficients $\hat{f}_e(k)$ may be defined, and the previous results hold also for $\tilde{J}_T = \mathrm{Tr}[\prod_{e=1}^{n} T_T(f_e)]$.*

Notes: 1) In the second case, the exact exponent of magnitude is unknown, except for the upper bound $\alpha\, d$.

2) The result above is a refinement of a limit theorem of Grenander and Szegö [46] concerning traces of products of truncated Toeplitz operators. Under the current strengthened integrability conditions, it was obtained when $d = 1$ in the discrete case in [14] and in the continuous case in [38]; as we show below, the result holds in fact in any dimension d (after modifying the necessary integrability condition for the asymptotic variance, in accordance to the Hölder-Young-Brascamp-Lieb inequality).

Following the proof of [12], we see that part 1 of Theorem 1.2 follows from Lemmas 0.1, 1.1 above, yielding the convergence of the normalized variance:

$$\lim_{T\to\infty} \frac{\chi_{2,T}}{T^d} = J(0, ..., 0) = \int_{\lambda \in S} f_1(\lambda) f_2(\lambda) ... f_k(\lambda)\mu(d\lambda).$$

Part 2 (see the proof of Theorem 3.1(c) below) implies that for $k \ge 3$, the cumulants satisfy

$$\lim_{T\to\infty} \frac{\chi_{k,T}}{T^{dk/2}} = 0,$$

and implies asymptotic normality. We arrive thus at the following multidimensional generalization of the results of Avram [10] and Ginovian [38].

Theorem 1.3. *Consider the quadratic functional*

$$Q_T = Q_T^{(1,1)} = \int_{t,s \in I_T} [X_t X_s - \mathbb{E} X_t \, X_s] \, \hat{b}(t-s) \nu(ds) \nu(dt),$$

where $X_t, \quad t \in I$ is a Gaussian random field with spectral density $f(\lambda) \in \mathbf{L}_p$. Assume the generating function b of the quadratic functional is such that $b(\lambda) \in \mathbf{L}_q$, that:

$$\frac{1}{p} + \frac{1}{q} \leq \frac{1}{2},$$

and that in the continuous case we have also $b(\lambda), f(\lambda) \in L_1$.

Then, the central limit theorem holds:

$$\lim T^{-d/2} Q_T \to N(0, \sigma^2), \quad T \to \infty,$$

where

$$\sigma^2 := 2(2\pi)^d \int_S b^2(\lambda) f^2(\lambda) d\lambda. \tag{1.8}$$

We present now one more result for Gaussian fields, which is related in the discrete case to the classical result of Breuer and Major [27], and in the continuous case to the result of Ivanov and Leonenko [53]. We note that these authors worked under time-domain assumptions, however, reasoning in the spectral domain with the methodology of [14] and the present paper – see Example 3.5 – immediately lead to the following result.

Theorem 1.4. *Let X_t, $t \in I$, be a Gaussian random field with spectral density $f(\lambda) \in \mathbf{L}_p$. Let $S_T = \int_{t \in I_T} P_l(X_t) \nu(dt)$, where $P_l(X_t)$ are univariate Appell (Hermite) polynomials and $l \geq 2$. Assume that:*

$$z := p^{-1} \leq 1 - \frac{1}{l}. \tag{1.9}$$

Then,

$$\sigma^2 := f^{(*,l)}(0) = \int_{y_1,\ldots,y_{l-1} \in S} f(y_1) f(y_2 - y_1) \ldots f(y_{l-1} - y_{l-2}) f(-y_{l-1}) \prod_{i=1}^{l-1} dy_i < \infty. \tag{1.10}$$

If, moreover, $\sigma \neq 0$, then the central limit theorem holds

$$\lim T^{-d/2} S_T \to N(0, \sigma^2).$$

Note: The difference between the integral representations of the variances (1.8) and (1.10) will be explained via graph theory in the next section.

2. Fejér graph/matroid integrals and graph/matroid convolutions

In this section, we introduce a unifying graph-theoretical framework for problems similar to those of the previous section.

Definition 2.1. *Let $G = (\mathcal{V}, \mathcal{E})$ denote a directed graph with V vertices, E edges, a basis of C independent cycles* [6] *and $co(G)$ components.* ***The incidence matrix*** *$M = \{M_{v,e}\}_{v\in\mathcal{V},e\in\mathcal{E}}$ of the graph is the $V \times E$ matrix with entries $[v,e] = \pm 1$ if the vertex v is the end/start point of the edge e, and 0 otherwise.*

A **circuit matrix** M^* is a $C \times E$ matrix whose rows are obtained by assigning arbitrary orientations to a basis of circuits (cycles) $c = 1, ..., C$ of the graph, and by writing each edge as a sum of $\pm$ the circuits it is included in, with the $\pm$ sign indicating a coincidence or opposition to the orientation of the cycle

Besides the graph framework, we will hint also to possible matroid generalizations. To clarify this point, let us start by quoting Tutte: "it is probably true that any theorem about graphs expressible in terms of edges and circuits exemplifies a more general result about vector matroids".

Let us recall briefly that matroids are a concept which formalizes the properties of the "rank function" $r(A)$ obtained by considering the rank of an arbitrary set of columns A in a given arbitrary matrix M. More precisely, a matroid is a pair $\mathcal{E}, r : 2^{\mathcal{E}} \to \mathbb{N}$ of a set $\mathcal{E}$ and a "rank like function" $r(A)$ defined on the subsets of $\mathcal{E}$. Matroids may also be defined in equivalent ways via their independent sets, via their bases (maximal independent sets), via their circuits (minimal dependent sets), via their spanning sets (sets containing a basis), or via their flats (sets which may not be augmented without increasing the rank). For precise definitions and for excellent expositions on graphs and matroids, see [61], [62] or [72].

The most familiar matroids, called *vectorial matroids*, are defined by the set $\mathcal{E}$ of columns of a matrix and by the rank function $r(A)$ which gives the rank of any set of columns A (matrices with the same rank function yield the same matroid).

Some useful facts from matroid theory are the fact that to each matroid M one may associate a **dual matroid**, with rank function

$$r^*(A) = |A| - r(M) + r(\mathcal{E} - A).$$

For vectorial matroids, the dual matroid is also vectorial, associated to any matrix whose rows span the space orthogonal to the rows of M. Furthermore, in the case of graphic matroids, the dual matroid is associated to the circuit matrix.

Tutte's "conjecture" holds true in our case: a matroid Szegö-type limit theorem was already given in [14]. However, for simplicity, we will restrict ourselves here to the particular case of **graphic matroids** associated to the incidence matrix M of an oriented graph. In this case, the proofs are more intuitive, due to the fact that the algebraic dependence structures translate into graph-theoretic concepts, like circuits corresponding to cycles, etc.

From here on, we will restrict ourselves to the graphic case, i.e. to the case when our dependence matrix is the incidence matrix of a directed graph.

Let $(S, d\mu)$ denote either $\mathbb{R}^d$ with Lebesgue measure, or the torus $[-\pi, \pi]^d$ with normalized Lebesgue measure, and let $f_e(\lambda) : S \to \mathbb{R},\ e = 1, \ldots, E$ denote a set of functions associated to the columns of M, which satisfy integrability conditions

$$f_e \in \mathbf{L}_{p_e}(S, d\mu),\ 1 \le p_e \le \infty. \tag{2.1}$$

[6] a basis of cycles is a set of cycles, none of which may be obtained via addition modulo 2 of other cycles, after ignoring the orientation

Let $\hat{f}_e(k), k \in I$, the Fourier transform of $f_e(\lambda)$:

$$\hat{f}_e(k) = \int_S e^{ik\lambda} f_e(\lambda)\mu(d\lambda), \quad k \in I,$$

where $I = \mathbb{Z}^d$ in the torus case and $I = \mathbb{R}^d$ in the case $S = \mathbb{R}^d$, respectively. In this last case, we would also need to assume that $f_e \in L_1(\mathbb{R}^d, d\mu)$, for the Fourier transform to be well defined. However, all our analytic results concern the spectral domain, and hence this assumption will not be necessary.

Our object of interest, in its "time domain representation", is:

$$\begin{aligned} \tilde{J}_T &= \tilde{J}_T(M, f_e, e = 1, ..., E) \\ &= \int_{j_1,...,j_V \in I_T} \hat{f}_1(i_1)\hat{f}_2(i_2)...\hat{f}_E(i_E) \prod_{v=1}^{V} \nu(dj_v), \end{aligned} \tag{2.2}$$

where $i = (i_1, ..., i_E) = (j_1, ..., j_V)M = jM$, where $\nu(dj_v)$ stands for Lebesgue measure and counting measure, respectively, and where in the torus case the linear combinations are computed modulo $[-\pi,\ \pi]^d$, so that the linear map $jM : S^V \to S^E$ is well defined.

Note. To keep the transparent analogy with the case, when $d = 1$, we make the following convention concerning notations. Here and in what follows let us treat a product of a vector, whose components are d-dimensional, and a matrix (or another vector) with scalar components in a specific sense: we will still perform multiplication component-wise according to the usual rule, and as a result we obtain a vector, whose components are d-dimensional again (or, correspondingly, just d-dimensional vector).

A **Fejér graph integral** is the expression obtained by replacing the sequences $\hat{f}_e(t)$ in (2.2) by their Fourier representations $\hat{f}_e(t) = \int_S f_e(\lambda)e^{it\lambda}\mu(d\lambda)$: under the assumption $f_e \in L_1(S, d\mu)$, an easy computation (see [12], Lemma 1) shows that (2.2) may be written also as the integral (2.4) below. We introduce however a more general concept.

Definition 2.2. *Let $(S, d\mu)$ denote either $\mathbb{R}^d$ with Lebesgue mesure, or the torus $[-\pi, \pi]^d$ with normalized Lebesgue mesure. Let M be a matrix of dimensions $V \times E$, with arbitrary coefficients in the first case and with integer coefficients in the second case. Let $f_e(\lambda) : S \to \mathbb{R},\ e = 1, \ldots, E$ denote a set of functions associated to the columns of M. Suppose these functions satisfy integrability conditions*

$$f_e \in \mathbf{L}_{p_e}(S, d\mu),\ 1 \le p_e \le \infty, \tag{2.3}$$

A Fejér matroid integral is defined by the following "spectral representation":

$$\begin{aligned} J_T &= J_T(M, f_e, e = 1, ..., E) \\ &= \int_{\lambda_1,...,\lambda_E \in S} f_1(\lambda_1)f_2(\lambda_2)...f_E(\lambda_E) \prod_{v=1}^{V} \Delta_T(u_v) \prod_{e=1}^{E} \mu(d\lambda_e) \end{aligned} \tag{2.4}$$

where $\Delta_T(u)$ is a kernel defined by (0.1), where $(u_1, ..., u_V)' = M(\lambda_1, ..., \lambda_E)'$, and where in the torus case the linear combinations are computed modulo $[-\pi,\ \pi]^d$.

A Fejér matroid integral will be called a Fejér graph integral for graphic matroids associated to the incidence matrix M of a directed graph G. In this case, the functions and kernels in (2.4) are associated respectively to the edges and vertices of the graph.

The cycle graph/Toeplitz example Consider the particular case of a cyclic graph with n edges. In this case, the matrix M with n columns and rows, is:

$$M = \begin{pmatrix} -1 & 0 & 0 & . & ... & 0 & 1 \\ 1 & -1 & 0 & 0 & ... & 0 & 0 \\ 0 & 1 & -1 & \ddots & ... & \vdots & \vdots \\ \vdots & & \ddots & \ddots & & & \\ 0 & 0 & & \ddots & & -1 & 0 \\ 0 & 0 & . & . & . & 1 & -1 \end{pmatrix}$$

and its Fejér graph integral is given by:

$$J_T = \int_{\lambda_1,\dots,\lambda_n \in S} f_1(\lambda_1) f_2(\lambda_2) \dots f_n(\lambda_n) \prod_{v=1}^{n} \Delta_T(\lambda_v - \lambda_{v+1}) \prod_{e=1}^{n} \mu(d\lambda_e).$$

Note. For analytical results concerning only Fejér matroid integrals as defined by (2.4), the condition $f_e \in L_1(S, d\mu)$ is unnecessary.

3. Limit theory for Fejér graph integrals

The main points of the limit theory for Fejér graph integrals, to be presented now, are that:

(1) Under certain Hölder-Young-Brascamp-Lieb conditions necessary to ensure the existence of the limiting integral, the following convergence holds as $T \to \infty$:

$$\boxed{T^{-d}\, J_T(M, f_1, \dots, f_E) \to \int_{S^C} f_1(\lambda_1) f_2(\lambda_2) \dots f_E(\lambda_E) \prod_{c=1}^{C} \mu(dy_c)}, \tag{3.1}$$

where $(\lambda_1, \dots \lambda_E) = (y_1, \dots, y_C)M^*$ (with every λ_e reduced modulo $[-\pi, \pi]^d$ in the torus case), M^* being any matrix whose rows span the space orthogonal to the rows of M, and C being the rank of M^*. Informally, the kernels disappear in the limit, giving rise to the "dual matroid" M^*.

(2) When the Hölder-Young-Brascamp-Lieb conditions do not hold, then, cf. part c) of the theorem below, the normalization defined in part a) will lead to a zero limit.

Theorem 3.1. *Suppose that $f_e \in L_{p_e}(d\mu)$ for part a), and $f_e \in \mathbf{L}_{p_e}(d\mu)$ for parts b),c) and set $z = (p_1^{-1}, \dots, p_E^{-1})$.*

Let $J_T = J_T(M, f_1, \dots, f_E)$ denote a Fejér matroid integral and let $r(A), r^*(A)$ denote respectively the ranks of a set of columns in M and in the dual matroid M^*.

Suppose that for every row l of the matrix M, one has $r(M) = r(M_l)$, where M_l is the matrix with the row l removed. Then:

a)

$$\boxed{J_T(M, f_1, \dots, f_E) \le c_M T^{d\, \alpha_M(z)}} \tag{3.2}$$

where c_M is a constant independent of z and

$$\begin{cases} \textit{in the discrete case} & \alpha_M(z) \textit{ is given by (3.8)} \\ \textit{in the continuous case} & \alpha_M(z) = \alpha_M^c(z) = co(M) + (\sum_e z_e - C)_+ \end{cases}$$

b) *If* $\alpha_M(z) = V - r(M) = co(M)$*, or, equivalently,*

$$\begin{array}{ll} \sum_{j\in A} z_j \le r^*(A), \quad \forall A & \textit{in the discrete case} \\ \sum_e z_e \le C & \textit{in the continuous case} \end{array} \tag{3.3}$$

then

$$\lim_{T\to\infty} \frac{J_T(M)}{T^{d\, co(M)}} = k_M\ \mathcal{J}(M^*, f_1, ..., f_E), \tag{3.4}$$

where

$$\boxed{\mathcal{J}(M^*, f_1, ..., f_E) = \int_{S^C} f_1(\lambda_1) f_2(\lambda_2)...f_E(\lambda_E) \prod_{c=1}^{C} \mu(dy_c)} \tag{3.5}$$

and where $(\lambda_1, ...\lambda_E) = (y_1, ..., y_C)M^*$ *(with every* λ_e *reduced modulo* $[-\pi, \pi]^d$ *in the discrete case), and* C *denotes the rank of the dual matroid* M^**.*

c) *If a strict inequality* $\alpha_M(z) > co(M)$ *holds, then the inequality (3.2) of Theorem 3.1 a) may be strengthened to:*

$$J_T(M) = o(T^{d\,\alpha_M(z)})$$

Remark: The results of this theorem, that is, the expression of $\alpha_M(z)$ and the limit integral $\mathcal{J}(M^*) = \mathcal{J}(M^*, f_1, ..., f_E)$, as well as the convergence conditions of integrals depend on M, M^* only via the two equivalent rank functions $r(A), r^*(A)$, i.e. only via the matroid dependence structure between the columns, and not on the chosen representing matrices.

Proof: The proof of part b) of Theorem 3.1 is essentially identical with that given in [12], up to the modification of the integrability conditions and the appearance of the extra constant k_M. For completeness, we sketch now this proof, for a connected graph (w.l.o.g.).

Note first that in a connected graph there are only $V-1$ independent rows of the incidence matrix M (or independent variables u_j), since the sum of all the rows is 0 (equivalently, $u_V = -\sum_{v=1}^{V-1} u_v$). Thus, $r(M) = V-1$, $co(M) = 1$, and the order of magnitude appearing in the normalization is just T^d.

The main idea behind the proof of Theorem 3.1 b) are a change of variables and applying the continuity of graph convolutions:

(1) **Change of variables.** Fix a basis $y_1, ..., y_C$ in the complement of the space generated by the u_v's, $v = 1, \dots, V$, switch to the variables $u_1, ..., u_{V-1}, y_1, ..., y_C$ and integrate in (2.4) first over the variables y_c's, $c = 1, \dots, C$. This is more convenient in the graphic case, since, after fixing an arbitrary spanning tree $\mathcal{T}$ in the graph, the complementary set of edges $\mathcal{T}^c$ furnishes a maximal set of independent cycles (with cardinality C). Assume w.l.o.g. that in the list $(\lambda_1, ..., \lambda_E)$, the edges in $\mathcal{T}^c$ are listed first, namely $(\lambda_e, e \in \mathcal{T}^c) = (\lambda_1, ..., \lambda_C)$. We make the change of variables $y_1 = \lambda_1, ..., y_C = \lambda_C$, and $(u_1, ..., u_{V-1})' = \tilde{M}(\lambda_1, ..., \lambda_E)'$, where $\tilde{M}$ denotes the first $V-1$ rows of the incidence matrix M. Thus,

$$(y_1, ..., y_C, u_1, ..., u_{V-1})' = \begin{pmatrix} I_C & 0 \\ \tilde{M}_C & \tilde{M}_V \end{pmatrix} (\lambda_1, ..., \lambda_E)'$$

where the first rows are given by an identity matrix I_C completed by zeroes and where $\tilde{M}_C, \tilde{M}_V$ denote the first C columns/ next $V-1-C$ columns of the matrix $\tilde{M}$.

Inverting the transformation above yields:

$$(\lambda_1, ..., \lambda_E) = (y_1, ..., y_C, u_1, ..., u_{V-1}) \begin{pmatrix} I_C & -\tilde{M}'_C \tilde{M}_V^{-1} \\ 0 & \tilde{M}_V^{-1} \end{pmatrix} = (y_1, ..., y_C, u_1, ..., u_{V-1}) \begin{pmatrix} M^* \\ \hline N \end{pmatrix}, \quad (3.6)$$

that is, it turns out that the first rows of the inverse matrix are precisely the dual matroid M^*.

Definition 3.2. The function

$$h_{M^*,N}(u_1, ..., u_{r(M)}) = \int_{y_1,...,y_C \in S} f_1(\lambda_1) f_2(\lambda_2) ... f_E(\lambda_E) \prod_{c=1}^{C} d\mu(y_c) \quad (3.7)$$

where λ_e are represented as linear combinations of $y_1, ..., y_C, u_1, ..., u_{V-1}$ via the linear transformation (3.6) will be called a **matroid/graph convolution** depending on whether the matroid is graphic or not.

The change to the variables $y_1, ..., y_C, u_1, ..., u_{V-1}$ and integration over $y_1, ..., y_C$ transforms the Fejér graph integral into the following integral of the product of a "graph convolution" and a Fejér kernel:

$$J_T(M) = \int_{u_1,...,u_{V-1} \in S} h_{M^*,N}(u_1, ..., u_{V-1}) \prod_{v=1}^{V} \Delta_T(u_v) \prod_{v=1}^{V-1} d\mu(u_v).$$

Recalling that the Fejér kernel converges under appropriate conditions to Lebesgue measure on the set $u_1 = ... = u_{V-1} = 0$, we find, just as in the cycle case, that part b) of Theorem 3.1 will be established once the convergence of the kernels and the continuity of the graph convolutions $h(u_1, ..., u_{r(M)})$ in the variables $(u_1, ..., u_{r(M)})$ is established.

(2)

Lemma 3.3. ***The continuity of graph convolutions.*** *The "graph convolution" function $C(u_1, ..., u_{n-1})$ defined in (1.5) is bounded and continuous if (1.2) holds with integrability indices satisfying the power counting condition (1.6).*

The proof is essentially the same as in the cycle case. Note that the function $h : \mathbb{R}^{V-1} \to \mathbb{R}$ is a composition

$$h_{M^*,N}(u_1, ..., u_{V-1}) = J(M^*, T_1(f_1), ..., T_E(f_E))$$

of the continuous functionals

$$T_e(u_1, ..., u_{V-1}) : \mathbb{R}^{V-1} \to \mathbf{L}_{p_e}$$

and of the functional

$$J(M^*, f_1, ..., f_E) : \prod_{e=1}^{E} \mathbf{L}_{p_e} \to \mathbb{R}.$$

The functional T_e is defined by $T_e(u_1,...,u_{V-1}) = f_e(\cdot + \sum_v u_v N_{v,e})$, where the $N_{v,e}$ are the components of the matrix N in (3.6). The functionals T_e are clearly continuous when f_e is a continuous function, and this continues to be true for functions $f_e \in \mathbf{L}_{p_e}$, since these can be approximated in the L_{p_e} sense by continuous functions. Thus, under our assumptions, the continuity of the functional $h_{M^*,N}(u_1,...,u_{V-1})$ follows automatically from that of $J(M^*, f_1,..., f_E)$.

Finally, under the "power counting conditions" (3.3), the continuity of the function $J(M^*, f_1,..., f_E)$ follows from the Hölder-Brascamp-Lieb-Barthe inequality:

$$|J(M^*, f_1,..., f_E)| \leq \prod_{e=1}^{E} \|f_e\|_{p_e}$$

(see Theorem C.1).

In conclusion, the convergence of the Fejér kernels to a δ measure implies the convergence of the scaled Fejér graph integral

$$J_T(M, f_e, e = 1..., E) \quad \text{to} \quad \mathcal{J}(M^*, f_e, e = 1..., E),$$

establishing Part b) of the theorem.

The proof of parts a), c) are postponed to section 3.2.

Remarks: 1) In the spatial statistics papers ([7], [9]), the continuity of the graph convolutions $h_{M^*,N}(u_1,...,u_{V-1})$ was assumed to hold, and indeed checking whether this assumption may be relaxed was one of the outstanding difficulties for the spatial extension.

2) It is not difficult to extend this approach to the case of several components and then to the matroid setup. In the first case, one would need to choose independent cycle and vertex variables $y_1,...,y_{r(M^*)}$ and $u_1,...,u_{r(M)}$, note the block structure of the matrices, with each block corresponding to a graph component, use the fact that for graphs with several components, the rank of the graphic matroid is $r(M) = V - co(G)$ and finally Euler's relation $E - V = C - co(G)$, which ensures that

$$E = \big(V - co(G)\big) + C = r(M) + r(M^*).$$

3) An important feature of the discrete case is that the limiting result (relation (3.4)) when f_e are complex exponentials is straightforward, implying therefore immediately theorem 3.1 in this case, by the multilinearity of $\mathcal{J}(M^*, f_1,..., f_E)$ and $T^{-d}J(M, f_1,..., f_E)$ and by Lemma 3.4 below.

Lemma 3.4. *For any matrix with $r(M) = V - 1$, any set of integers $b = (b_e, e = 1,..., E)$, and any functions $f_e(\lambda_e) = e^{i\lambda_e b_e}, e = 1,..., E$, theorem 3.1 holds, i.e.:*

$$\lim_{T\to\infty} \frac{\int_{S^E} e^{i<b,\lambda>} \prod_{v=1}^{V} \Delta_T(u_v) \prod_{e=1}^{E} \mu(d\lambda_e)}{T^d} = k_M \int_{S^C} e^{i<\lambda,b>} \prod_{c=1}^{C} \mu(dy_c)$$

Proof: Aside from the constant k_M, the stated RHS (limiting value) above is:

$$\int_{S^C} e^{i<\lambda,b>} \prod_{c=1}^{C} \mu(dy_c) = \int_{S^C} e^{i<y,M^*b>} \prod_{c=1}^{C} \mu(dy_c) = 1_{M^*b=0} = 1_{b\in R(M)}$$

where $R(M)$ denotes the subspace generated by the rows of M.

Now the LHS in (3.4), before scaling, is:

$$\int_{S^E} e^{i<b,\lambda>} \int_{I_T^V} e^{-i<sM,\lambda>} \prod_{v=1}^{V} \nu(ds_v) \prod_{e=1}^{E} \mu(d\lambda_e) = \int_{I_T^V} \prod_{v=1}^{V} \nu(ds_v) (\int_{S^E} e^{i<b-sM,\lambda>} \prod_{e=1}^{E} \mu(d\lambda_e))$$
$$= \int_{I_T^V} \prod_{v=1}^{V} \nu(ds_v) 1_{b-sM=0} = \nu(s : sM = b, s \in I_T)$$

Now for a matrix with integer entries it holds that:

$$E(T) := \nu(s : sM = b, s \in I_T)$$

is either

$$E(T) \begin{cases} = 0 & \text{if } b \notin R(M), \\ \sim T^D \mu(I_1 \cap ker(M)) & \text{if } b \in R(M), \end{cases}$$

the second statement being tantamount to the definition of the Lebesgue measure. Thus, the result holds with $k_M = \mu(I_1 \cap ker(M))$. See for more details [14].

Note also that $E(T)$ is an "Ehrhart quasi polynomial", whose next coefficients are related to other geometric characteristics of I_1, which should allow developing correction terms to Theorem 3.1.

Note: The lemma above may be interpreted as saying that the measures on $(S)^E$ given by the "multiple Fejér kernels"

$$T^{-d} \Delta_T(-\sum_{v=1}^{V-1} \lambda_v) \prod_{v=1}^{V-1} (\Delta_T(\lambda_v) \nu(d\lambda_\nu))$$

converge weakly as $T \to \infty$ to the uniform measure on the subspace $\lambda = yM^*$ (since the Fourier coefficients converge).

3.1. The upper bound for the order of magnitude of Fejér matroid integrals

We turn now to the "upper bound exponent" $\alpha_M(z)$ for the order of magnitude of Fejér matroid integrals (useful when it is not precisely d). In the discrete case [14], the exponent $\alpha_M(z)$ of this upper bound turns out to be d times the solution of a graph optimization problem:

$$\boxed{\alpha_M(z) = co(M) + \max_{A \subset 1,\ldots,E} [\sum_{j \in A} z_j - r^*(A)]} \tag{3.8}$$

or equivalently,

$$\boxed{\alpha_M(z) = \max_{A \subset 1,\ldots,E} [co(M-A) - \sum_{j \in A} (1 - z_j)]} \tag{3.9}$$

where $co(M-A)$ represents the number of remaining components, after the edges in A have been removed, and, for a general Fejér graph integral, we define

$$co(M-A) = V - r(M-A). \tag{3.10}$$

Note that for connected graphs and under the power counting conditions $\sum_{j \in A} z_j \leq r^*(A)$, this exponent reduces to d, as in Theorem 1.2.

We will call the problem (3.9) a **graph breaking problem: find a set of edges whose removal maximizes the difference between the number of remaining components and** $\sum_{j\in A}(1-z_j)$, or, equivalently, the difference between $\sum_{j\in A} z_j$ and the dual rank $r^*(A)$.

Note: In the following examples an important role is played by the **"maximal breaking"** $A = M$ and **"no breaking** $A = \emptyset$**"** sets A, which yield often the solution of the optimal breaking problem. It is useful to introduce therefore the lower bound:

$$\begin{aligned}\alpha^c_M(z) &= \max_{A\in\{\mathcal{E},\emptyset\}}[co(M-A) - \sum_{j\in A}(1-z_j)] = \max\{co(M), \sum_e (z_e-1)+V\}\\ &= \max\{co(M), co(M)+\sum_e z_e - C\} = co(M) + (\sum_e z_e - C)_+ \ ,\end{aligned} \tag{3.11}$$

where the equality before the last holds by Euler's relation $C = (E-V)_+$ (C denotes the number of cycles, and, more generally, the rank of the dual matroid M^*).

Note that in the case of a cycle graph of size m, this reduces to

$$\alpha^c_m(z) = \max\{1, \sum_{e=1}^m z_e\},$$

as stated in Theorem 1.2, and that the expression (3.11) turns out to yield the upper bound exponent in the continuous case.

Note: Theorem 3.1 c) may be used for establishing convergence to 0 of higher order cumulants, whenever these may be written as sums of Fejér graph integrals, by computing bounds of the form

$$T^{d\alpha_k(z)}$$

for Fejér graph integrals intervening in the cumulants of order k. Since the typical CLT normalization is $T^{d/2}$, it will suffice then establishing **"cumulant inequalities"**

$$d\alpha_k(z) < kd/2 \Leftrightarrow \boxed{\alpha_k(z) < k/2}$$

where $\alpha_k(z)$ is the exponent appearing in the expansion of the k-th cumulant. In fact, this may be strengthened (cf. Theorem 3.1) to include z satisfying the equality $\alpha_k(z) = k/2$, if $k \geq 3$.

In conclusion, establishing normality is reduced to computing the functions $\alpha_k(z), k \geq 3$, i.e., to solving a sequence of graph breaking problems.

Example 3.5. The general structure of the intervening graphs for the k'th cumulant of sums S_T of the m'th Appell polynomial of a Gaussian sequence is provided by graphs belonging to the set $\Gamma(m,k)$ of all connected graphs with no loops over k vertices, each of degree m (see [14]). Let $z = p^{-1}$ denote the integrability exponent of the spectral density.

The cumulant inequality corresponding to the **"maximal breaking=MB"** of the k'th cumulant graph, which typically yields a facet of the power counting polytope (PCP) (at least for $k = 2$) is given in this example by:

$$\alpha_k(z) = \sum_e (z_e - 1) + V = \frac{km}{2}(z-1) + k \leq \frac{k}{2} \Leftrightarrow \frac{1}{m} \leq 1 - z$$

At the limiting point $1-z=\frac{1}{m}$, the cumulant exponents are $\alpha_k(z)=\frac{k}{2}$, ensuring negligibility for $k\geq 3$.

The discrete and continuous case may be unified here by asking for integrability at $z=1-\frac{1}{m}$, since in the discrete case the extension to smaller values of z is trivial.

Example 3.6. A similar analysis holds in the case of cumulants of quadratic forms in Appell polynomials $P_{m,n}(X_t,X_s)$. Note that while the number of graphs intervening increases considerably, the number of extremal points of the PCP is just 4 – see Figure 2.

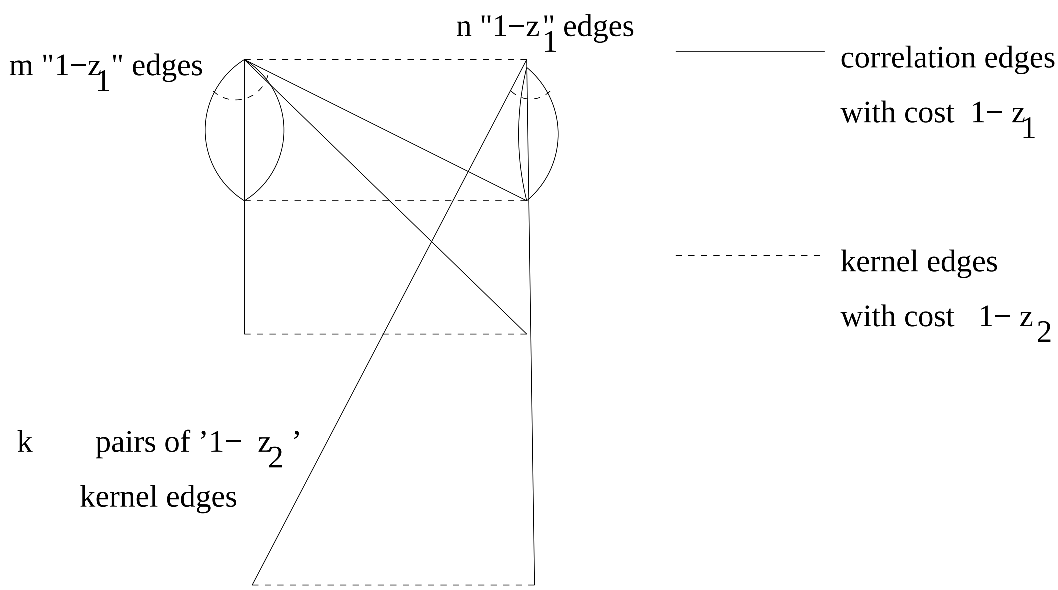

FIGURE 1. **The graphs appearing in the expansion of cumulants of quadratic forms. Here k=4, m=5, n=4. The figure displays only some of the k(m+n)/2=18 correlation edges.**

The graphs intervening are – see Figure 1 – all the graphs belonging to the set $\Gamma(m,n,k)$ of all connected bipartite graphs with no loops whose vertex set consists of k pairs of vertices. The "left" vertex of each pair arises out of the first m terms : $X_{t_1},...,X_{t_m}$: in the diagram formula, and the "right" vertex of each pair arises out of the last n terms : $X_{s_1},...,X_{s_n}$: The edge set consists of:

(1) k "kernel edges" pairing each left vertex with a right vertex. The kernel edges will contribute below terms involving the function $b(\lambda)$.
(2) A set of "correlation edges", always connecting vertices in different rows, and contributing below terms involving the function $f(\lambda)$). They are arranged such that each left vertex connects to m and each right vertex connects to n such edges, yielding a total of $k(m+n)/2$ correlation edges.

Thus, the k "left vertices" are of degree $m+1$, and the other k vertices are of degree $n+1$. (The "costs $1-z_1, 1-z_2$" mentioned in Figure 1 refer to (3.8)).

The PCP domain in the discrete case (which is precisely the convergence domain of the integrals defining the limiting variance), is indicated below, when $m<n$, in terms of the integrability indices $z=(z_1,z_2)$ of f and b (i.e. $f\in L_{p_1}, b\in L_{p_2}$).

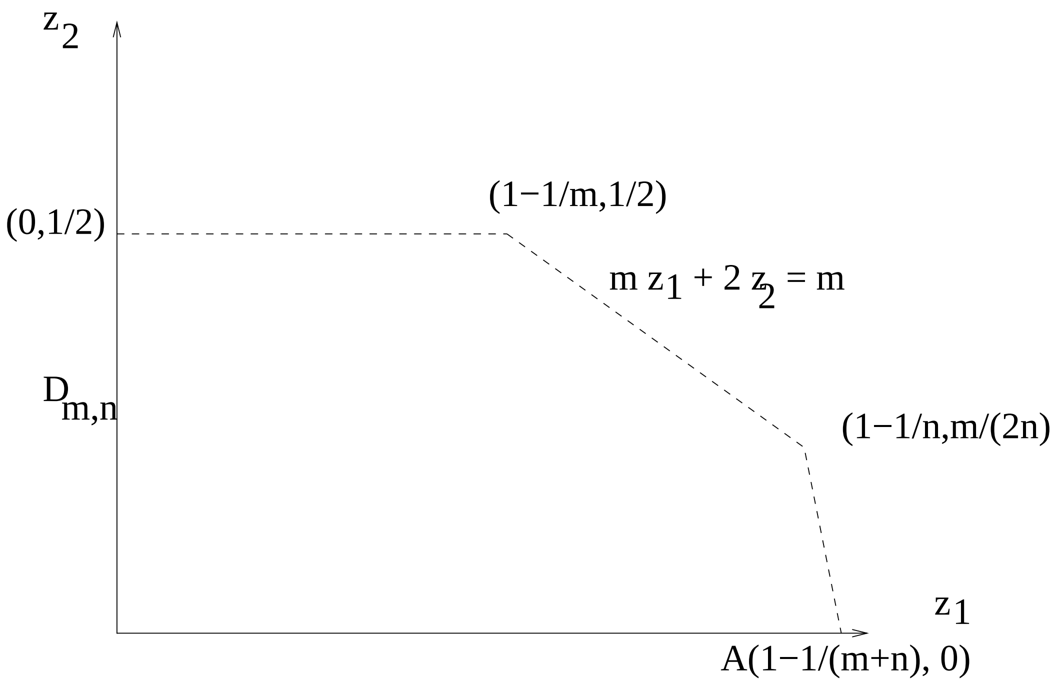

Figure 2. **The domain of the central limit theorem, discrete case**

When $m < n$, there are only three segments on the undominated boundary of the PCP, connecting respectively the extremal points (A, B), (B, C) and (C, D)(with coordinates $A(1 - 1/(m+n), 0)$, $B(1-1/n, m/(2n))$, $C(1-1/m, 1/2)$, $D(0, 1/2)$), and correspond respectively to the breakings indicated below:

$$\begin{cases} \text{total breaking} & 2k - \frac{k(m+n)}{2}(1-z_1) - k(1-z_2) \le \frac{k}{2} \\ & \Leftrightarrow \frac{3}{2} \le \frac{(m+n)}{2}(1-z_1) + (1-z_2) \\ \text{breaking all } z_2 \text{ and the left } z_1 \text{ edges} & \max_k \frac{k+1}{k} - m(1-z_1) - (1-z_2) \le \frac{1}{2} \\ & \Leftrightarrow 1 \le m(1-z_1) + (1-z_2) \\ \text{breaking all } z_2 \text{ edges} & \max_k \frac{2}{k} - (1-z_2) =\le \frac{1}{2} \\ & \Leftrightarrow \frac{1}{2} \le z_2 \end{cases}$$

In the continuous case, the domain is just the lower segment between the points A and B in figure 2.

3.2. **Proof of Theorem 3.1, parts a), c)**

We turn now to theorem 3.1 a), c), generalizing Theorem 2 and Corollary 1 of [14].

For part a), let us apply the Hölder-Young-Brascamp-Lieb inequality with optimally chosen integrability parameters s_v^{-1}:

$$|J_T(M, f_1, ..., f_E)| \leq K \prod_{v=1}^{V} \|\Delta_T\|_{s_v^{-1}} \prod_{e=1}^{E} \|f_e\|_{z_e^{-1}} \tag{3.12}$$

$$\leq K' T^{d \sum_{v=1}^{V}(1-s_v^{-1})} \prod_{e=1}^{E} \|f_e\|_{z_e^{-1}} \tag{3.13}$$

under the constraint that $(s_1, ..., s_V, z_1, ..., z_E)$ satisfy the power counting conditions, and where we used the kernel estimate

$$\|\Delta_T(\lambda)\|_{s^{-1}} \leq C_s T^{d(1-s)}, \forall s \in [0, 1)$$

(see Appendix B).

The optimization problem for s_v in the discrete case:

$$\min_{s_1,...,s_V} d \sum_{v=1}^{V}(1 - s_v^{-1}), \text{ where } (s_1, ..., s_V, z_1, ..., z_E) \in \text{ PCP},$$

has the same constraints as Lemma 2 in [14], except that the objective is multiplied by d. Hence, in the torus case, the exponent is simply d times the one dimensional exponent of Theorem 1, [14].

In the continuous case $S = \mathbb{R}^d$, we note first that when $\sum_e z_e \leq C$ the result follows, just as Theorem 3.1 b), from the Hölder-Young-Brascamp-Lieb inequality, while in the other case $\sum_e z_e \geq C$, the extra constraint $\sum_v s_v = E - \sum_e z_e$ yields the one-dimensional exponent as $\alpha_M(z) = co(M) + \sum_e z_e - C$.

For part c), we approximate our functions by continuous, bounded functions of bounded support, for which conditions (3.3) of part b) hold. For these approximants, it follows from the convergence $\frac{J_T(M)}{T^{d\, co(M)}} \to k_M\, \mathcal{J}(M^*, f_1, ..., f_E)$ that $J_T(M) = o(T^{d\, co(M)+a})$, $\forall a > 0$. The result follows then for the functions from $\mathbf{L}_p$ spaces, by the definition of these spaces. Note that in the discrete case, the same argument was applied based on trigonometric polynomials – see [14], proof of Corollary 1.

4. Applications

4.1. Central limit theorems for bilinear forms of moving averages

We assume below that our stationary random field $X_t,\ t \in I$ admits a representation as a linear/moving averages random field.

For discrete parameter it means that

$$X_t = \sum_{u \in \mathbb{Z}^d} \hat{a}(t-u)\xi(u), \ \sum_{u \in \mathbb{Z}^d} \hat{a}^2(u) < \infty, \ t \in \mathbb{Z}^d, \tag{4.1}$$

where $\xi(u)$, $u \in \mathbb{Z}^d$, are independent random variables indexed by $\mathbb{Z}^d$ with $\mathbb{E}\xi(0) = 0$ and such that $E\,|\xi(0)|^k \leq c_k < \infty,\ k = 1, 2, \ldots$. In this case

$$c_k(t_1, ..., t_k) = cum_k\{X_{t_1}, ..., X_{t_k}\} = d_k \sum_{s \in \mathbb{Z}^d} \prod_{j=1}^{k} \hat{a}(t_j - s), \tag{4.2}$$

where d_k is the k'th cumulant of $\xi(0)$.

For continuous parameter we assume that

$$X_t = \int_{u \in \mathbb{R}^d} \hat{a}(t-u)\xi(du), \quad t \in \mathbb{R}^d, \tag{4.3}$$

with a square-integrable kernel $\hat{a}(t)$, $t \in I$, with respect to a independently scattered random measure with finite second moment, that is a homogeneous random measure $\xi(A),\ A \subset \mathbb{R}^d$, with finite second moments and independent values over disjoint sets (see, for instance, Rajput and Rosinski [65] or Kwapien and Woyczynski [55]). That is, for each Borel A, $\xi(A)$ is an infinitely divisible random variable whose cumulant function can be written as

$$\kappa(z) = \log Ee^{iz\xi(A)} = izm_0(A) - \frac{1}{2}z^2 m_1(A) + \int_{\mathbb{R}} \left(e^{izx} - 1 - iz\tau(x)\right) Q(A, dx), \tag{4.4}$$

where m_0 is a signed measure, m_1 is a positive measure, $Q(A, dx)$ (for fixed A) is a measure on $\mathbb{R}^1$ without atoms at 0, such that $\int_{\mathbb{R}} \min\left\{1, |x|^2\right\} Q(A, dx) < \infty$, and where $\tau(x) = x$ if $|x| \leq 1$, and $\tau(x) = x/|x|$, if $|x| > 1$.

For example, if $I = \mathbb{R}$, then $\xi(A)$ is a set indexed Lévy process with finite second moments and stationary intensity proportional to the Lebesgue measure.

We also assume that Q factorizes as $Q(A, dx) = M(A)W(dx)$, where $M(A)$ is a $\sigma-$finite measure, and W is some Lévy measure on $\mathbb{R}^1$, such that for some $\varepsilon > 0$ and $\lambda > 0$

$$\int_{(-\varepsilon,\varepsilon)} e^{\lambda u} W(du) < \infty.$$

This implies that

$$\int_{\mathbb{R}} |u|^k W(du) < \infty,\ k \geq 2,$$

and that the cumulant function $\kappa(z)$ is analytical in a neighborhood of 0.

Necessary and sufficient conditions of existence of the integral (as limit in probability of integrals of simple functions)

$$\int_A f(s) d\xi(s)$$

can be found in [65]. Note that, for $d = 1$, an integrals become integrals with respect to Lévy process $L(t),\ t \in \mathbb{R}^1$, and $\kappa(z) = \log Ee^{izL(1)}$.

For Lebesgue measures $m_0,\ m_1$ and Q, one can prove (by using product integration) that

$$\log E \exp\left\{i(z_1 X_{t_1} + \cdots + z_k X_{t_k}\right\} = \int_{\mathbb{R}^d} \kappa\left(\sum_{j=1}^{k} z_j \hat{a}(t_j - s)\right) ds \tag{4.5}$$

if $\hat{a} \in L_1 \cap L_2$. From (4.5) if can be seen that random field (4.3) is homogeneous in a strict sense.

We assume that $m_1 = 0$, that is, $\mathbb{E}\xi(I_1) = 0$, then the last formula holds for $\hat{a} \in L_2$.

We obtain that

$$c_k(t_1, ..., t_k) = cum_k\{X_{t_1}, ..., X_{t_k}\} = d_k \int_{\mathbb{R}^d} \prod_{j=1}^{k} \hat{a}(t_j - s)ds, \tag{4.6}$$

where d_k is the k'th cumulant of $\xi(I_1)$ with I_1 being the unit rectangle, that is $d_k = \kappa^k(0)/i^k,\ k \geq 2$.

We assume from now on that $\mathbb{E}\xi(I_1) = 0$, and use the same notation for both discrete and continuous cases

$$X_t = \int_{u \in I} \hat{a}(t-u)\xi(du), \quad t \in I, \tag{4.7}$$

where $I = \mathbb{Z}^d$ in the discrete case and $I = \mathbb{R}^d$ in the continuous case.

For various conditions which ensure that (4.7) is well-defined, see, for example, Anh, Heyde and Leonenko [6], p. 733, and references therein.

Note: By choosing an appropriate "Green function" $\hat{a}(t)$, this very general class of processes includes the solutions of many interesting differential equations with random noise $\xi(du)$, like, for example, generalized Ornstein-Uhlenbeck processes in $\mathbb{R}$ [6].

We will assume here that all moments for our stationary field X_t exist.

The advantage of the linear representation assumption (4.7) and (4.4), (4.5) is the explicit representation of cumulants – see for example Theorem 2.1 of [6]:

$$c_k(t_1, ..., t_k) = d_k \int_{s \in I} \prod_{j=1}^{k} \hat{a}(t_j - s)\ \nu(ds), \tag{4.8}$$

where d_k is the k'th cumulant of $\xi(I_1)$ with I_1 being the unit rectangle.

In the spectral domain, we get

$$f_k(\lambda_1, ..., \lambda_{k-1}) = d_k\ a(-\sum_{i=1}^{k-1} \lambda_i) \prod_{i=1}^{k-1} a(\lambda_i) = \prod_{i=1}^{k} a(\lambda_i)\delta(\sum_{j=1}^{k} \lambda_j). \tag{4.9}$$

For $k = 2$, we will denote the spectral density by $f(\lambda) = f_2(\lambda) = d_2 a(\lambda)a(-\lambda)$.

We can formulate now a central limit theorems for quadratic functional of a linear field, which is a generalization of the results of Giraitis and Surgailis [44] and Giraitis and Taqqu [42] (see also references therein) for discrete time processes. This next theorem follows from the results of Sections 3 and 4 (the proof is almost identical to the proof of Theorem 4 of Avram [14], the expression for the variance can be obtained by direct computations).

Theorem 4.1. *Let $X_t = \int_{u \in\ I} \hat{a}(t-u)\xi(\ du), \quad t \in\ I$, be a linear random field with a square integrable kernel $\hat{a}(t), t \in I$, and a random measure $\xi(du)$ admitting all moments and let*

$$Q_T = Q_T^{(1,1)} = \int_{t,s \in I_T} [X_t X_s - \mathbb{E}X_t\, X_s]\, \hat{b}(t-s)\nu(ds)\nu(dt).$$

We assume that $f(\lambda) = (2\pi)^d\, |a(\lambda)|^2 \in \mathbf{L}_p$ and $b(\lambda) \in \mathbf{L}_q$, and in the continuous case we assume also that $b(\lambda) \in \mathbf{L}_q \cap L_1$.

Assume that:

$$\frac{1}{p}+\frac{1}{q}\leq\frac{1}{2}.$$

Then, the central limit theorem holds:

$$\lim T^{-d/2}Q_T\to N(0,\sigma^2),\quad T\to\infty,$$

where

$$\sigma^2:=2(2\pi)^d d_2^2\int_S b^2(\lambda)f^2(\lambda)d\lambda+(2\pi)^d d_4\left(\int_S b(\lambda)f(\lambda)d\lambda\right)^2$$

where d_k is the k'th cumulant of $\xi(I_1)$ with I_1 being the unit rectangle, that is
$d_2=\mathbb{E}\xi(I_1)^2, d_4=\mathbb{E}(\xi(I_1)^4)-2[\mathbb{E}(\xi(I_1)^2)]^2$ *in the continuous case and d_k is the k'th cumulant of $\xi(0)$ in the discrete case, that is $d_2=\mathbb{E}\xi(0)^2$, $d_4=\mathbb{E}(\xi(0)^4)-2[\mathbb{E}(\xi(0)^2)]^2$.*

4.2. **Minimum contrast estimation based on the Whittle contrast function**

The class of Whittle estimators is the most popular in applications (see Whittle [74], [75], Giraitis and Surgailis [41], Fox and Taqqu [33], Heyde and Gay [47], [48], Heyde [49], Gao, Anh and Heyde [37], Leonenko and Sakhno [57], see also the references therein).

In what follows we will consider continuous time linear processes ($d=1$) whose spectral densities of all orders exist and admit the representation of the form (4.9).

We begin with the following assumption.

A.I. Let X_t, $t\in I_T=\left[-\frac{T}{2},\frac{T}{2}\right]$, be an observation of a real-valued measurable stationary linear process X_t, $t\in\mathbb{R}^1$, with zero mean and the family of spectral densities (4.9). Let $a(\lambda)=a\left(\lambda;\theta^{(1)}\right), d_k= d_k\left(\theta^{(2)}\right)$, that is, $f_2(\lambda)=f(\lambda,\theta)$, $\lambda\in\mathbb{R}^1$, $\theta=\left(\theta^{(1)},\theta^{(2)}\right)$, $\theta\in\Theta\subset\mathbb{R}^m$, where Θ is a compact set, and the true value of the parameter $\theta_0\in int\Theta$, the interior of Θ. Suppose further that $f(\lambda;\theta_1)\not\equiv f(\lambda;\theta_2)$ for $\theta_1\neq\theta_2$, almost everywhere in $\mathbb{R}^1$ with respect to the Lebesgue measure.

Consider the Whittle contrast process (or objective function)

$$U_T(\theta)=\frac{1}{4\pi}\int_{\mathbb{R}^1}\left(\log f(\lambda;\theta)+\frac{I_T(\lambda)}{f(\lambda;\theta)}\right)w(\lambda)\,d\lambda,\tag{4.10}$$

where $I_T(\lambda)$ is the periodogram of the second order

$$I_T(\lambda)=\frac{1}{2\pi T}\left|\int_{I_T}X_te^{-it\lambda}dt\right|^2,\quad\lambda\in\mathbb{R}^1,\tag{4.11}$$

and $w(\lambda)$ is a symmetric about $\lambda=0$ function such that all considered integrals are well defined and which will satisfy some conditions given below; in some cases we can choose $w(\lambda)=\frac{1}{1+\lambda^2}$.

Introduce the Whittle contrast function

$$K(\theta_0;\theta)=\frac{1}{4\pi}\int_{\mathbb{R}^1}\left(\frac{f(\lambda;\theta_0)}{f(\lambda;\theta)}-1-\log\frac{f(\lambda;\theta_0)}{f(\lambda;\theta)}\right)w(\lambda)\,d\lambda.\tag{4.12}$$

To state the result on consistency of the minimum contrast estimator based on the contrast process (4.10) we will need the following conditions on the spectral density $f(\lambda;\theta)$ and the weight function $w(\lambda)$.

A.II. $f(\lambda;\theta_0)\,w(\lambda)\,\frac{1}{f(\lambda;\theta)}\in L_1\left(\mathbb{R}^1\right)\cap L_2\left(\mathbb{R}^1\right),\quad\forall\theta\in\Theta.$

A.III. There exists a function $v(\lambda), \lambda \in \mathbb{R}^1$, such that
(i) the function $h(\lambda;\theta) = v(\lambda)\frac{1}{f(\lambda;\theta)}$ is uniformly continuous in $\mathbb{R}^1 \times \Theta$;
(ii) $f(\lambda;\theta_0)\frac{w(\lambda)}{v(\lambda)} \in L_1\left(\mathbb{R}^1\right) \cap L_2\left(\mathbb{R}^1\right)$.

Theorem 4.2. *Let the assumptions A.I to A.III be satisfied. Then the function $K(\theta_0;\theta)$ defined by (4.12) is the contrast function for the contrast process $U_T(\theta)$ defined by (4.10). The minimum contrast estimator $\widehat{\theta}_T$ defined as*

$$\widehat{\theta}_T = \arg\min_{\theta\in\Theta} U_T(\theta) \tag{4.13}$$

is a consistent estimator of the parameter θ, that is, $\widehat{\theta}_T \to \theta_0$ in P_0-probability as $T \to \infty$.

The above theorem can be obtained as a consequence of a more general result by Leonenko and Sakhno [57] (Theorem 3), one needs just to rewrite for the case of linear processes the corresponding conditions on spectral densities, which become of much simpler form.

Next set of assumptions (in addition to the above ones) is needed to state the result on asymptotic normality of the estimator (4.13).

A.IV. The function $\frac{1}{f(\lambda;\theta)}$ is twice differentiable in a neighborhood of the point θ_0 and
(i) $f(\lambda;\theta_0)\, w(\lambda)\frac{\partial^2}{\partial\theta_i\partial\theta_j}\frac{1}{f(\lambda;\theta)} \in L_1\left(\mathbb{R}^1\right) \cap L_2\left(\mathbb{R}^1\right)$, $i,j = 1,...,m$, $\theta \in \Theta$;
(ii) $f(\lambda;\theta_0) \in \mathbf{L}_p\left(\mathbb{R}^1\right)$, $w(\lambda)\frac{\partial}{\partial\theta_i}\frac{1}{f(\lambda;\theta)} \in \mathbf{L}_q\left(\mathbb{R}^1\right)$,
for some p, q such that $\frac{1}{p}+\frac{1}{q} \le \frac{1}{2}$, $i = 1,...,m$, $\theta \in \Theta$;
(iii) $T^{1/2}\int_{\mathbb{R}^1}(EI_T(\lambda) - f(\lambda;\theta_0))w(\lambda)\frac{\partial}{\partial\theta_i}\frac{1}{f(\lambda;\theta)}\,d\lambda \to 0$ as $T \to \infty$,
for all $i = 1,...,m, \theta \in \Theta$;
(iv) the second order derivatives $\frac{\partial^2}{\partial\theta_i\partial\theta_j}\frac{1}{f(\lambda;\theta)}, i = 1,...,m$, are continuous in θ.

A.V. The matrices $W_1(\theta) = \left(w_{ij}^{(1)}(\theta)\right)_{i,j=1,\dots,m}$, $W_2(\theta) = \left(w_{ij}^{(2)}(\theta)\right)_{i,j=1,\dots,m}$, $V(\theta) = (v_{ij}(\theta))_{i,j=1,\dots,m}$ are positive definite, where

$$w_{ij}^{(1)}(\theta) = \frac{1}{4\pi}\int_{\mathbb{R}^1} w(\lambda)\frac{\partial}{\partial\theta_i}\log f(\lambda;\theta)\frac{\partial}{\partial\theta_j}\log f(\lambda;\theta)\,d\lambda, \tag{4.14}$$

$$w_{ij}^{(2)}(\theta) = \frac{1}{4\pi}\int_{\mathbb{R}^1} w^2(\lambda)\frac{\partial}{\partial\theta_i}\log f(\lambda;\theta)\frac{\partial}{\partial\theta_j}\log f(\lambda;\theta)\,d\lambda. \tag{4.15}$$

$$v_{ij}(\theta) = \frac{1}{8\pi}\frac{d_4}{d_2^2}\int_{\mathbb{R}^1} w(\lambda)\frac{\partial}{\partial\theta_i}\log f(\lambda;\theta)\,d\lambda\int_{\mathbb{R}^1} w(\lambda)\frac{\partial}{\partial\theta_j}\log f(\lambda;\theta)\,d\lambda. \tag{4.16}$$

Theorem 4.3. *Let the assumptions A.I to A.V be satisfied. Then as $T \to \infty$*

$$T^{1/2}\left(\widehat{\theta_T} - \theta_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}_m\left(0, W_1^{-1}(\theta_0)\,(W_2(\theta_0) + V(\theta_0))W_1^{-1}(\theta_0)\right),$$

where $N_m(\cdot,\cdot)$ denotes the m-dimensional Gaussian law.

Reasonings for the proof of Theorems 4.2, 4.3 are given in the next section.

Comparing the above theorem with a more general result stated in [57], one can see that the set of conditions for the case of linear processes becomes of much simpler form, but the most important improvement is in condition A.IV(ii), which was achieved due to the application of the Theorem 4.1 (see the proof). Note that corresponding condition for the case of general processes, formulated in [57], unfortunately, is difficult to check in general situation.

Remark 4.4. Condition A.IV(iii) will hold, e.g., if $f(\lambda;\theta)$ is differentiable with respect to λ and

$$\int_{\mathbb{R}^1} f'_\lambda(\lambda;\theta_0))w(\lambda)\frac{\partial}{\partial\theta_i}\frac{1}{f(\lambda;\theta)}d\lambda<\infty,$$

or under any conditions which assure

$$\int_{\mathbb{R}^1} |f(\lambda+h;\theta_0)-f(\lambda;\theta_0)|w(\lambda)\frac{\partial}{\partial\theta_i}\frac{1}{f(\lambda;\theta)}d\lambda\le C|h|^a,$$

for $a>\frac{1}{2}$ and C being a constant.

Example. Estimation of fractional Riesz-Bessel motion (FRBM) (see Appendix A for details and definition of FRBM in non-Gaussian case). Let X_t, $t\in\mathbb{R}^1$, be a non-Gaussian Riesz-Bessel stationary motion, that is a stationary linear process with the spectral density of the form

$$f(\lambda)=f(\lambda,\theta)=\frac{c}{|\lambda|^{2\alpha}(1+\lambda^2)^\gamma},\quad \lambda\in\mathbb{R}^1, \tag{4.17}$$

where the unknown vector parameter $\theta=(\gamma,\alpha,c)'\in\Theta$, Θ being a compact subset of $\left[\frac{1}{2},\infty\right)\times\left(0,\frac{1}{2}\right)\times(0,\infty)$. Note that the index α determines the long-range dependence of FRBM, and the parameter γ is another fractal index connected to Hausdorff dimension of paths of the stochastic process. Note that procedure of discretazation leads to the loss of information of one parameter γ, which is important for applications in both turbulence and finance theory. That is why a direct method of estimation of both parameters from continuous data looks appropriate.

For this model we can choose the weight function $w(\lambda)=\frac{1}{1+\lambda^2}$, $\lambda\in\mathbb{R}^1$, to satisfy the conditions needed for consistency of the estimator (4.13), that is, for Theorem 4.2 to hold. However, to satisfy all the conditions needed for Theorem 4.3 we choose the weight function $w(\lambda)=\frac{\lambda^{2b}}{(1+\lambda^2)^a}$, $\lambda\in\mathbb{R}^1$, where a and b satisfy the restrictions:$\{b>1\}\wedge\{a>b+2\}\wedge\{a>A+2\}$,where we have denoted by A the length of the finite interval carrying the admissible values of the parameter γ. With such a choice of the weight function we have the convergence

$$T^{1/2}\left(\widehat{\theta_T}-\theta_0\right)\xrightarrow{\mathcal{D}}N_3\left(0,W_1^{-1}(\theta_0)(W_2(\theta_0)+V(\theta_0))W_1^{-1}(\theta_0)\right)\quad\text{as}\quad T\to\infty,$$

where the elements of the matrices W_1 and W_2 are of the following form:

$$\begin{aligned}
w_{11}^{(1\vee2)} &= \frac{1}{4\pi}\int_{\mathbb{R}^1} w^{1\vee2}(\lambda)\left(\ln\left(1+\lambda^2\right)\right)^2 d\lambda;\\
w_{22}^{(1\vee2)} &= \frac{1}{4\pi}\int_{\mathbb{R}^1} w^{1\vee2}(\lambda)\left(\ln\left(\lambda^2\right)\right)^2 d\lambda;\\
w_{33}^{(1\vee2)} &= \frac{1}{4\pi}c_0^{-2}\int_{\mathbb{R}^1} w^{1\vee2}(\lambda)\,d\lambda;
\end{aligned}$$

$$
\begin{aligned}
w_{12}^{(1\vee 2)} &= w_{21}^{(1\vee 2)} = \frac{1}{4\pi}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\ln\left(1+\lambda^2\right)\ln\left(\lambda^2\right)d\lambda;\\
w_{13}^{(1\vee 2)} &= w_{31}^{(1\vee 2)} = -\frac{1}{4\pi}c_0^{-1}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\ln\left(1+\lambda^2\right)d\lambda;\\
w_{23}^{(1\vee 2)} &= w_{32}^{(1\vee 2)} = -\frac{1}{4\pi}c_0^{-1}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\ln\left(\lambda^2\right)d\lambda;
\end{aligned}
$$

$$
\begin{aligned}
v_{11} &= \frac{1}{8\pi}\frac{d_4}{d_2^2}\left(\int_{\mathbb{R}^1} w(\lambda)\ln\left(1+\lambda^2\right)d\lambda\right)^2;\\
v_{22} &= \frac{1}{8\pi}\frac{d_4}{d_2^2}\left(\int_{\mathbb{R}^1} w(\lambda)\ln\left(\lambda^2\right)d\lambda\right)^2;\\
v_{33} &= \frac{1}{8\pi}\frac{d_4}{d_2^2}c_0^{-2}\left(\int_{\mathbb{R}^1} w(\lambda)\,d\lambda\right)^2;
\end{aligned}
$$

$$
\begin{aligned}
v_{12} &= v_{21} = \frac{1}{8\pi}\frac{d_4}{d_2^2}\int_{\mathbb{R}^1} w(\lambda)\ln\left(1+\lambda^2\right)d\lambda\int_{\mathbb{R}^1} w(\lambda)\ln\left(\lambda^2\right)d\lambda;\\
v_{13} &= v_{31} = -\frac{1}{8\pi}\frac{d_4}{d_2^2}c_0^{-1}\int_{\mathbb{R}^1} w(\lambda)\ln\left(1+\lambda^2\right)d\lambda\int_{\mathbb{R}^1} w(\lambda)\,d\lambda;\\
v_{23} &= v_{32} = -\frac{1}{8\pi}\frac{d_4}{d_2^2}c_0^{-1}\int_{\mathbb{R}^1} w(\lambda)\ln\left(\lambda^2\right)d\lambda\int_{\mathbb{R}^1} w(\lambda)\,d\lambda.
\end{aligned}
$$

In the above formulae we mean that the weight function $w(\lambda)$ is involved to the expressions for $w_{ij}^{(1)}$ in the 1st power and to the expressions for $w_{ij}^{(2)}$ in the 2d power. From the above formulae we see that the covariance matrix of the limiting normal law has the charming feature that it appears not depending on the values α_0 and γ_0.

Remark 4.5. Continuous version of Gauss-Whittle objective function with the weight function $w(\lambda) = \frac{1}{1+\lambda^2}$ had been used in [37] for the estimation of the Gaussian processes in stationary and nonstationary cases respectively.

4.3. **Minimum contrast estimation based on the Ibragimov contrast function**

We consider now the minimum contrast functional motivated by the paper of Ibragimov [52], see also Anh, Leonenko and Sakhno [7].

We assume condition **A.I** and introduce the following condition

B. I. There exists a nonnegative function $w(\lambda), \lambda \in \mathbb{R}$, such that

(i) $w(\lambda)$ is symmetric about $\lambda = 0 : w(\lambda) = w(-\lambda)$;

(ii) $w(\lambda) f(\lambda;\theta)$ is in $L_1(\mathbb{R})$ for $\forall\theta \in \Theta$.

Under the condition B.I, we set

$$
\sigma^2(\theta) = \int_{\mathbb{R}} f(\lambda;\theta)\,w(\lambda)\,d\lambda
$$

and consider the factorization of the spectral density

$$f(\lambda;\theta)=\sigma^2(\theta)\,\psi(\lambda;\theta)\,,\quad \lambda\in\mathbb{R},\theta\in\Theta.$$

For the function $\psi(\lambda,\theta)\,,\lambda\in\mathbb{R},\ \theta\in\Theta$, we have

$$\int_{\mathbb{R}}\psi(\lambda;\theta)\,w(\lambda)\,d\lambda=1$$

and we additionally suppose

B. II. The derivatives $\nabla_\theta\psi(\lambda;\theta)$ exist and

$$\nabla_\theta\int_{\mathbb{R}}\psi(\lambda;\theta)\,w(\lambda)\,d\lambda=\int_{\mathbb{R}}\nabla_\theta\psi(\lambda;\theta)\,w(\lambda)\,d\lambda=0,$$

that is we can differentiate under the integral sign in the above integral.

Consider the following contrast process (or objective function):

$$U_T(\theta)=-\int_{\mathbb{R}}I_T(\lambda)\,w(\lambda)\log\psi(\lambda;\theta)\,d\lambda,\quad \theta\in\Theta. \tag{4.18}$$

Define also the function

$$K(\theta_0;\theta)=\int_{\mathbb{R}}f(\lambda;\theta_0)\,w(\lambda)\log\frac{\psi(\lambda;\theta_0)}{\psi(\lambda;\theta)}d\lambda,\quad \theta_0,\theta\in\Theta. \tag{4.19}$$

B. III. $f(\lambda;\theta_0)\,w(\lambda)\log\psi(\lambda;\theta)\in L_1\left(\mathbb{R}^1\right)\cap L_2\left(\mathbb{R}^1\right),\quad \forall\theta\in\Theta.$

B. IV. There exists a function $v(\lambda)\,,\lambda\in\mathbb{R}^1$, such that

(i) the function $h(\lambda;\theta)=v(\lambda)\log\psi(\lambda;\theta)$ is uniformly continuous in $\mathbb{R}^1\times\Theta$;

(ii) $f(\lambda;\theta_0)\frac{w(\lambda)}{v(\lambda)}\in L_1\left(\mathbb{R}^1\right)\cap L_2\left(\mathbb{R}^1\right).$

Theorem 4.6. *Let conditions AI, B.I - B.IV be satisfied. Then the function $K(\theta_0;\theta)$ defined by (4.19) is the contrast function for the contrast process $U_T(\theta)$ defined by (4.18). Moreover the minimum contrast estimator $\widehat{\theta}_T$ defined as*

$$\widehat{\theta}_T=\underset{\theta\in\Theta}{\arg\min}U_T(\theta)\,, \tag{4.20}$$

is a consistent estimator of the parameter θ, that is, $\widehat{\theta}_T\to\theta_0$ in P_0-probability as $T\to\infty$, and the estimator

$$\widehat{\sigma}^2_T=\int_{\mathbb{R}^n}I_T(\lambda)\,w(\lambda)\,d\lambda$$

is a consistent estimator of the parameter $\sigma^2(\theta)$, that is, $\widehat{\sigma}^2_T\to\sigma^2(\theta_0)$ in P_0-probability as $T\to\infty$.

To formulate the result on the asymptotic distribution of the minimum contrast estimator (4.20) we need some further conditions.

B. V. The function $\psi(\lambda;\theta)$ is twice differentiable in a neighborhood of the point θ_0 and

(i) $f(\lambda;\theta)\,w(\lambda)\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\psi(\lambda,\theta)\in L_1(\mathbb{R})\cap L_2(\mathbb{R})\,,\ i,j=1,...,m,\ \theta\in\Theta;$

(ii) $f(\lambda;\theta_0)\in\mathbf{L}_p\left(\mathbb{R}^1\right),\ w(\lambda)\frac{\partial}{\partial\theta_i}\log\psi(\lambda,\theta)\in\mathbf{L}_q\left(\mathbb{R}^1\right),$

for some p, q such that $\frac{1}{p}+\frac{1}{q}\le\frac{1}{2}$, $i=1,...,m,\ \theta\in\Theta;$

(iii) $T^{1/2}\int_{\mathbb{R}^1}(EI_T(\lambda)-f(\lambda;\theta_0))w(\lambda)\frac{\partial}{\partial\theta_i}\log\psi(\lambda;\theta)\ d\lambda\to 0$ as $T\to\infty$, for all $i=1,...,m,\theta\in\Theta$;

(iv) the second order derivatives $\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\psi(\lambda;\theta), i=1,...,m$, are continuous in θ.

B. VI. The matrices $S(\theta)=(s_{ij}(\theta))_{i,j=1,\dots,m}$ and $A(\theta)=(a_{ij}(\theta))_{i,j=1,\dots,m}$ are positive definite where

$$s_{ij}(\theta)=\int_{\mathbb{R}}f(\lambda;\theta)\,w(\lambda)\,\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\psi(\lambda;\theta)\,d\lambda$$

$$=\sigma^2(\theta)\int_{\mathbb{R}}w(\lambda)\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\psi(\lambda,\theta)-\frac{1}{\psi(\lambda,\theta)}\frac{\partial}{\partial\theta_i}\psi(\lambda,\theta)\,\frac{\partial}{\partial\theta_j}\psi(\lambda,\theta)\right]d\lambda,$$

$$a_{ij}(\theta)=4\pi\int_{\mathbb{R}^1}f^2(\lambda;\theta)\,w^2(\lambda)\,\frac{\partial}{\partial\theta_i}\log\psi(\lambda;\theta)\,\frac{\partial}{\partial\theta_j}\log\psi(\lambda;\theta)\,d\lambda$$

$$+2\pi\frac{d_4}{d_2^2}\int_{\mathbb{R}}\frac{w(\lambda)\,f(\lambda;\theta)}{\psi(\lambda;\theta)}\frac{\partial}{\partial\theta_i}\psi(\lambda;\theta)\,d\lambda\int_{\mathbb{R}}\frac{w(\lambda)\,f(\lambda;\theta)}{\psi(\lambda;\theta)}\frac{\partial}{\partial\theta_j}\psi(\lambda;\theta)\,d\lambda$$

$$=\ 4\pi\left(\sigma^2(\theta)\right)^2\int_{\mathbb{R}}w^2(\lambda)\,\frac{\partial}{\partial\theta_i}\psi(\lambda;\theta)\,\frac{\partial}{\partial\theta_j}\psi(\lambda;\theta)\,d\lambda$$

$$+2\pi\frac{d_4}{d_2^2}\left(\sigma^2(\theta)\right)^2\int_{\mathbb{R}}w(\lambda)\,\frac{\partial}{\partial\theta_i}\psi(\lambda;\theta)\,d\lambda\int_{\mathbb{R}}w(\lambda)\,\frac{\partial}{\partial\theta_j}\psi(\lambda;\theta)\,d\lambda$$

Theorem 4.7. *Let the conditions AI, B.I - B.VI be satisfied. Then as* $T\longrightarrow\infty$

$$T^{1/2}\left(\widehat{\theta_T}-\theta_0\right)\xrightarrow{\mathcal{D}}N_m\left(0,S^{-1}(\theta_0)\,A(\theta_0)\,S^{-1}(\theta_0)\right),$$

where $N_m(\cdot,\cdot)$ *denotes the m-dimensional Gaussian law.*

Proofs of Theorems 4.2, 4.3, 4.6, 4.7.

The results on consistency of estimators (Theorems 4.2 and 4.6) are consequences of corresponding theorems stated for the general case in [57] for the Whittle functional and in [7], [8] for the case of Ibragimov functional. We present here reasonings for the proofs of Theorems 4.3 and 4.7, which make use of CLT for bilinear forms (Theorem 4.1 above). For the proofs the standard arguments based on Taylor's formula for $\nabla_\theta U_T\left(\widehat{\theta}_T\right)$ are used. Namely, we can write the relation

$$\nabla_\theta U_T\left(\widehat{\theta}_T\right)=\nabla_\theta U_T(\theta_0)+\nabla_\theta\nabla'_\theta U_T(\theta_T^*)\left(\widehat{\theta}_T-\theta_0\right),$$

where $|\theta_T^*-\theta_0|<\left|\widehat{\theta}_T-\theta_0\right|$.

It follows from the definition of minimum contrast estimators that for sufficiently large T

$$\nabla_\theta U_T(\theta_0)=-\nabla_\theta\nabla'_\theta U_T(\theta_T^*)\left(\widehat{\theta}_T-\theta_0\right),$$

therefore, to state the asymptotic normality for the estimator $\widehat{\theta}_T$, by Slutsky's arguments, one needs to deduce: (1) limit in probability for $\nabla_\theta\nabla'_\theta U_T(\theta_T^*)$ and (2) limiting normal law for $T^{1/2}\nabla_\theta U_T(\theta_0)$.

For the 1-st task we can use the same arguments as in the mentioned above papers, and to rewrite (simplify) corresponding conditions for the case of linear processes.

However, for the step (2) we can appeal now to Theorem 4.1. We provide the details below.

Consider firstly the case of Whittle functional. Limit in P_0-probability for $\nabla_\theta \nabla'_\theta U_T(\theta^*_T)$ is given by the matrix $W_1(\theta_0)$.

Next, consider

$$\nabla_\theta U_T(\theta_0) = \frac{1}{4\pi} \int_{\mathbb{R}} \left(\nabla_\theta \log f(\lambda;\theta)|_{\theta=\theta_0} + \nabla_\theta \left(\frac{1}{f(\lambda;\theta)} \right) \bigg|_{\theta=\theta_0} I_T(\lambda) \right) w(\lambda)\, d\lambda,$$

which can be written in the form

$$\begin{aligned} \nabla_\theta U_T(\theta_0) &= (J_T(\varphi_i) - J(\varphi_i))_{i=1,\dots,m} \\ &= \left(\int_{\mathbb{R}} \varphi_i(\lambda;\theta_0) I_T(\lambda)\, d\lambda - \int_{\mathbb{R}} \varphi_i(\lambda;\theta_0) f(\lambda;\theta_0)\, d\lambda \right)_{i=1,\dots,m} \end{aligned}$$

where

$$\varphi_i = \varphi_i(\lambda;\theta_0) = -\frac{1}{4\pi} \frac{1}{f^2(\lambda;\theta_0)} w(\lambda) \left(\frac{\partial}{\partial \theta_i} f(\lambda;\theta) \right) \bigg|_{\theta=\theta_0}, \quad i = 1, \dots, m.$$

Under the assumptions of Theorem 4.3 (see A.IV(ii)) in view of Theorem 4.1 we have the convergence

$$T^{1/2} (J_T(\varphi_i) - EJ_T(\varphi_i))_{i=1,\dots,m} \xrightarrow{\mathcal{D}} N_m(0, W_2(\theta_0) + V(\theta_0)), \tag{4.21}$$

where the the matrices $W_2(\theta_0)$ and $V(\theta_0)$ are defined in the assumption A.V.

Further, in view of the assumption A.IV(iii)

$$T^{1/2} (EJ_T(\varphi_i) - J(\varphi_i)) \to 0, \text{ as } T \to \infty$$

which, combined with (4.21), implies

$$T^{1/2} (J_T(\varphi_i) - J(\varphi_i))_{i=1,\dots,m} = T^{1/2} \nabla_\theta U_T(\theta_0) \xrightarrow{\mathcal{D}} N_m(0, W_2(\theta_0) + V(\theta_0)).$$

The case of Ibragimov functional is treated analogously. We have that $\nabla_\theta \nabla'_\theta U_T(\theta^*_T)$ converges in P_0-probability to the matrix $S(\theta_0)$. Further,

$$\nabla_\theta U_T(\theta_0) = -\int_{\mathbb{R}} I_T(\lambda)\, \nabla_\theta \log \psi(\lambda;\theta)|_{\theta=\theta_0} w(\lambda)\, d\lambda.$$

In view of B.II

$$\int_{\mathbb{R}} f(\lambda;\theta_0)\, \nabla_\theta \log \psi(\lambda;\theta)|_{\theta=\theta_0} w(\lambda)\, d\lambda = 0,$$

and we can write

$$\begin{aligned} \nabla_\theta U_T(\theta_0) &= (J_T(\varphi_i) - J(\varphi_i))_{i=1,\dots,m} \\ &= \left(\int_{\mathbb{R}} \varphi_i(\lambda;\theta_0) I_T(\lambda)\, d\lambda - \int_{\mathbb{R}} \varphi_i(\lambda;\theta_0) f(\lambda;\theta_0)\, d\lambda \right)_{i=1,\dots,m}, \end{aligned}$$

where now

$$\varphi_i = \varphi_i(\lambda;\theta_0) = w(\lambda)\,\frac{\partial}{\partial\theta_i}\log\psi(\lambda;\theta)\bigg|_{\theta=\theta_0}, \quad i=1,...,m.$$

Again in view of Theorem 4.1, under the assumption B.V(ii), we obtain the convergence

$$T^{1/2}\left(J(\varphi_i) - EJ_T(\varphi_i)\right) \xrightarrow{\mathcal{D}} N_m\left(0, A\left(\theta_0\right)\right), \text{ as } T\to\infty,$$

where the matrix $A(\theta_0)$ is defined in B.VI. By assumption B.V(iii) the convergence $T^{1/2}\nabla_\theta U_T(\theta_0) \xrightarrow{\mathcal{D}} N_m(0, A(\theta_0))$ follows.

Appendix A. Fractional Riesz-Bessel Motion

In this Appendix we mainly review a number results discussed in Gay and Heyde [36], Anh, Angulo and Ruiz-Medina [4], Anh, Leonenko and Mc Vinish [5], Anh and Leonenko [2], [3], Kelbert, Leonenko and Ruiz-Medina [54] (see also references therein). Also we introduce a not necessarily Gaussian Riesz-Bessel stationary process and formulate the central limit theorem for such a processes as well as for quadratic forms of such a processes.

The fractional operators are natural mathematical objects to describe the singular phenomena of random fields such as long range dependence or/and intermittency.

In particular Gay and Heyde [36] introduced a class of random fields as solutions of fractional Helmholtz equation driven by white noise, contained the fractional operator $(cI-\Delta)^{\alpha/2},\ c\geq 0$ (and its limit as $c\to 0 \quad (-\Delta)^{\alpha/2}$), where Δ is the d-dimensional Laplacian and I is the identity operator (see also [54] for properties of such fields and possible generalization). Anh, Angulo, Ruiz-Medina [4] (see also [5], [2], [3] and references therein) generalized the fractional stochastic equation of Gay and Heyde in which the fractional Helmgoltz operator $(cI-\Delta)^{\alpha/2},\ c\geq 0$ or the d-dimensional Laplacian $(c\to 0)$ is replaced by a fractional Laplace-type operator of the form $-(I-\Delta)^{\gamma/2}(-\Delta)^{\alpha/2}, \alpha>0,\ \gamma\geq 0$, where the operators $-(I-\Delta)^{\gamma/2}, \gamma\geq 0$, and $(-\Delta)^{\alpha/2},\ \alpha>0$, are interpreted as inverses to the Bessel and Riesz potentials (see [68], pp. 134-138), that is integral operators, whose kernels have a Fourier transforms $(2\pi)^{-d/2}\left(1+\|\lambda\|^2\right)^{-\gamma/2},\ \lambda\in\mathbb{R}^d$, and $(2\pi)^{-d/2}\|\lambda\|^{-\alpha},\ \lambda\in\mathbb{R}^d$, respectively. Then there exists a generalized random field $\zeta(x),\ x\in\mathbb{R}^d$, on fractional Sobolev space, which is defined by the equation

$$(I-\Delta)^{\gamma/2}(-\Delta)^{\alpha/2}\zeta(x) = e(x),\ x\in\mathbb{R}^d, \tag{A.1}$$

where $\left\{e(x),\ x\in\mathbb{R}^d\right\}$ is a Gaussian white noise or equivalently (in the sense of second-order moments) there exists a random field with the spectral density

$$f(\lambda) = \frac{c}{\|\lambda\|^{2\alpha}\left(1+\|\lambda\|^2\right)^{\gamma}},\ \lambda\in\mathbb{R}^d,\ c>0. \tag{A.2}$$

These random fields were named the fractional Riesz-Bessel motion.

For the random fields with stationary increments we assume $\alpha\in\left(\frac{d}{2},\frac{d}{2}+1\right), \gamma\geq 0$. In particular, for $d=1$, there exists a Gaussian stochastic process with stationary increments and the spectral density (A.2), where $\alpha\in\left(\frac{1}{2},\frac{3}{2}\right),\ \gamma\geq 0$. This fractional Riesz-Bessel motion (FRBM) is a generalization of the fractional Brownian motion (FBM) (see, for instance, Samorodnitsky and Taqqu [66]). FBM is a limiting case of the Riesz-Bessel (non-stationary) motion with $\gamma=0$ (in terms of the Hurst parameter

$H \in (0,1)$, the spectral density of the FBM with long-range dependence ($H \in (\frac{1}{2},0)$) is equal to $\frac{1}{|\lambda|^{2H+1}}$). The FRBM is not self-similar (unless when $\gamma = 0$), but it is locally self-similar.

For $d \geq 1$ the presence of the Bessel operator is essential for a study of homogeneous (and isotropic) solutions of (A.1) with spectral density (A.2), which requires $0 \leq \alpha < \frac{d}{2},\ \alpha + \gamma > \frac{d}{2}$; that is the condition $\gamma > 0$ is necessary for $f(\lambda) \in L_1(\mathbb{R}^d)$. Thus the homogeneous isotropic FRBM can be defined as a Gaussian random field with zero mean and covariance function of the form

$$B_{\alpha,\gamma}(x) = \int_{\mathbb{R}^d} e^{i<\lambda,x>} \frac{c}{\|\lambda\|^{2\alpha}\left(1+\|\lambda\|^2\right)^{\gamma}} d\lambda,\ x \in \mathbb{R}^d, \tag{A.3}$$

where $0 \leq \alpha < \frac{d}{2},\ \alpha + \gamma > \frac{d}{2}$. Note that for $\alpha = 0$, the covariance structure (A.3) belongs the Matérn class, that is with

$$c = \frac{\Gamma(\gamma)}{\pi^{d/2} 2^{d-1} \Gamma\left(\frac{2\gamma-d}{2}\right)};$$

$$B_{\alpha,\gamma}(x) = \frac{1}{2^{\frac{2-d}{2}} \Gamma\left(\frac{2\gamma-d}{2}\right)} \frac{K_{\frac{2\gamma-d}{2}}}{\|x\|^{\frac{d-2\gamma}{2}}},\ x \in \mathbb{R}^d,\ \gamma > \frac{d}{2}, \tag{A.4}$$

where

$$K_\nu(z) = \frac{1}{2} \int_0^\infty s^{\nu-1} \exp\left\{-\frac{1}{2}\left(s + \frac{1}{s}\right) z\right\} ds,\ z \geq 0,\ \nu \in R,$$

is the modified Bessel function of the third kind of order ν or Mc Donald's function. Note that

$$K_\nu(z) = K_{-\nu}(z), \quad K_\nu(z) \sim \Gamma(\nu) 2^{\nu-1} z^{-\nu},$$

for $\nu > 0$ as $z \to 0$,

$$K_{\frac{1}{2}}(z) = \sqrt{\frac{\pi}{2}} \frac{e^{-z}}{\sqrt{z}}.$$

Thus, we have $B_{\alpha,\gamma}(0) = 1$.

Note that for $d = 1,\ \alpha = 0,\ \gamma = 1$, the covariance structure (A.4) becomes $B_{0,1}(x) = e^{-x},\ x \geq 0$, that is stationary Gaussian Riesz-Bessel motion is identical to the Gaussian Ornstein-Uhlenbeck process.

Remark A.1. These results can be generalized to the case when the above fractional operator is replaced by the operator

$$H = \frac{\partial^\beta}{\partial t^\beta} + \mu \left(I - \Delta\right)^{\gamma/2} \left(-\Delta\right)^{\alpha/2},\ 0 \leq \beta \leq 2,\ \alpha > 0,\ \gamma \geq 0,$$

where $\frac{\partial^\beta}{\partial t^\beta}$ is the regularized fractional derivative. In particular, the Green function of the fractional heat equation: $Hu(t,x) = 0$, $t > 0$, $x \in \mathbb{R}^d$, can be given as inverse Fourier transform of the function

$$E_\beta\left(-\mu t^\beta \|\lambda\|^\alpha (1 + \|\lambda\|^2)^{\gamma/2}\right),\ t > 0,\ \lambda \in \mathbb{R}^d,$$

where

$$E_\beta(z) = \sum_{j=1}^{\infty} \frac{z^j}{\Gamma(\beta_j + 1)}, \ z \in \mathtt{C}, \ \beta > 0$$

is the Mittag-Leffler function (see [2], [3] for details and references).

In order to introduce a Riesz-Bessel motion driven by Lévy noise, we restrict our attention to the stationary case and $d = 1$ (replacing the space parameter x into t). For the function

$$a(\lambda) = \frac{\sqrt{c}}{(i\lambda)^\alpha (1 + i\lambda)^\gamma}, \ \lambda \in \mathbb{R}, \ a + \gamma > 1, \ \alpha \geq 0,$$

we introduce the function

$$\hat{a}(t) = \int_{\mathbb{R}} e^{it\lambda} a(\lambda) d\lambda = \begin{cases} \frac{2\pi}{\Gamma(\alpha+\gamma)} t^{\alpha+\gamma-1} e^{-1} \, {}_1F_1(\gamma, \alpha + \gamma; t), & t \geq 0, \\ 0, & t < 0, \end{cases} , \tag{A.5}$$

where the confluent hypergeometric function

$${}_1F_1(a, b; z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}, \ (c)_n = c(c+1) \cdots (c + n - 1), \ (c)_0 = 1.$$

The Riesz-Bessel motion driven by Lévy noise can be defined as the linear process

$$X_t = \int_{\mathbb{R}} \hat{a}(t - s) d\xi(s), \tag{A.6}$$

where $\xi(t), \ t \in R$ is a Lévy process with cumulant function

$$\kappa(z) = \log E \exp \left\{ iz\xi(1) \right\},$$

such that $\kappa^{(k)}(0) \neq 0, \ k \geq 2$, and $\hat{a}(.)$ is defined by (A.5). The k-th order spectral densities of the Riesz-Bessel motion driven by Lévy noise (A.6) take the form:

$$f_k(\lambda_1, \ldots, \lambda_{k-1}) = (2\pi)^{-k+1} i^{-k} \kappa^{(k)}(0) a(\lambda_1) \cdots a(\lambda_{k-1}) \overline{a(\lambda_1 + \cdots + \lambda_{k-1})}, \tag{A.7}$$

which reduces to the second-order spectral density

$$\begin{aligned} f_2(\lambda) &= \frac{c}{|\lambda|^{2\alpha} (1 + \lambda^2)^\gamma}, \qquad \text{(A.8)} \\ c &= \frac{\kappa^{(2)}(0)}{2\pi}, \ 0 \leq \alpha < \frac{1}{2}, \ \alpha + \gamma > \frac{1}{2}, \ \lambda \in \mathbb{R}^1. \end{aligned}$$

For the Gaussian case, of course, $\kappa^{(k)}(0) = 0, \ k \geq 3$.

Note that for $\alpha = 0, \ \gamma = 1$ we arrive to the Ornstein-Uhlenbeck process driven by Lévy noise ([6]).

As a consequence of the Theorems of Sections 4 and 5 we obtain the following result for the linear process (A.6). (Cf. also with Theorem 4.1)

Theorem A.2. *Consider the Riesz-Bessel stationary motion (A.6) and assume that all cumulants of Lévy process are finite. Let*

$$S_T = \int_{-T/2}^{T/2} X_s ds, \qquad Q_T = \int_{-T/2}^{T/2}\int_{-T/2}^{T/2} \hat{b}(t-s)\left[X_t X_s - E X_t X_s\right] dtds,$$

then:
i) if

$$\alpha + \gamma > \frac{1}{2}, \quad \alpha \le 0,$$

then the central limit theorem holds:

$$T^{-1/2} S_T \to N(0, \sigma^2), T \to \infty,$$

where

$$\sigma^2 = \kappa^{(2)}(0);$$

ii) if for some $p > 1, q > 1$, such that $\frac{1}{p} + \frac{1}{q} \le \frac{1}{2}$, we have

$$b(\lambda) \in L_q, \alpha + \gamma > \frac{1}{2p}, \quad \alpha < \frac{1}{2p},$$

then the central limit theorem holds:

$$T^{-1/2} Q_T \to N(0, \sigma^2), T \to \infty,$$

where

$$\sigma^2 = 2\kappa^{(2)}(0)\int_{\mathbb{R}} b^2(\lambda)\frac{1}{|\lambda|^{4\alpha}(1+\lambda^2)^{2\gamma}}d\lambda + \kappa^{(4)}(0)\frac{1}{(2\pi)^2}\int_{\mathbb{R}} b(\lambda)\frac{1}{|\lambda|^{2\alpha}(1+\lambda^2)^{\gamma}}d\lambda.$$

Appendix B. Kernel estimates

Consider the Dirichlet type kernel

$$\Delta_T(\lambda) = \int_{t \in I_T} e^{it\lambda}\nu(dt).$$

When $d = 1$ and $I_1 = [-1/2, 1/2]$ ($I_T = I_1 T$), one gets the clasical discrete/continuous time Dirichlet kernels:

$$\Delta_T(\lambda) = \sum_{-T/2}^{T/2} e^{it\lambda} = \frac{\sin((T+1)\lambda/2)}{\sin(\lambda/2)}, \quad \Delta_T(\lambda) = \int_{-T/2}^{T/2} e^{it\lambda}dt = \frac{\sin(T\lambda/2)}{\lambda/2},$$

respectively. For general d and $I_T = [-T/2, T/2]^d$, putting $\lambda = (\lambda_1, .., \lambda_d)$, it follows that $\Delta_T(\lambda) = \prod_{j=1}^{d} \Delta_T(\lambda_j)$.

Note that in the continuous case, by scaling, one finds

$$\|\Delta_T(\lambda)\|_p = T^{1-1/p} C_p, \quad \lambda \in \mathbb{R}, \quad 1 < p < \infty, \tag{B.1}$$

$$\|\Delta_T(\lambda)\|_p = T^{d(1-1/p)} C_p^d, \quad \lambda \in \mathbb{R}^d, \quad 1 < p < \infty. \tag{B.2}$$

with $C_p = (2\int_R |\frac{\sin(z)}{z}|^p dz)^{\frac{1}{p}}$.

In the discrete case, similar estimates may be obtained by using the inequality

$$\left|\frac{\sin((T+1)\lambda/2)}{\sin(\lambda/2)}\right| \le \tilde{C}\frac{T}{1+T|\lambda|}, \lambda \in [-\pi, \pi).$$

We find then:

$$\|\Delta_T(\lambda)\|_p \le T_p^{1-1/p}\tilde{C}^{\frac{1}{p}}\left(\int\limits_{\mathbb{R}} \frac{dz}{(1+|z|^p)}\right)^{\frac{1}{p}}, \quad \lambda \in [-\pi, \pi), 1 < p < \infty,$$

$$\|\Delta_T(\lambda)\|_p \le T_p^{d(1-1/p)}\left[\tilde{C}^{\frac{1}{p}}\left(\int\limits_{\mathbb{R}} \frac{dz}{(1+|z|^p)}\right)^{\frac{1}{p}}\right]^d, \quad \lambda = (\lambda_1, .., \lambda_d) \in [-\pi, \pi)^d, 1 < p < \infty.$$

In the case of the Euclidean ball $I_T = B_T = \{t \in \mathbb{R}^d : \|t\| \le T/2\}$, we find again by scaling in the continuous case

$$\Delta_T(\lambda) = \int_{B_T} e^{it\lambda}dt = (2\pi)^{\frac{d}{2}} J_{d/2}(\|\lambda\| \frac{T}{2}) / \|\lambda\|^{d/2}, \quad \lambda \in \mathbb{R}^d,$$

where $J_\nu(z)$ is the Bessel function of the first kind and order ν. It is known that $J_\nu(z) \le const/\sqrt{z}$ for a large z, thus for the ball

$$\|\Delta_T(\lambda)\|_p = \begin{cases} T^{(1-\frac{1}{p})}C_p, & d = 1, \quad p > 1, \\ T^{d(\frac{1}{2}-\frac{1}{p})}\bar{C}_p, & d \ge 2, \quad p > \frac{2d}{d+1}, \end{cases}$$

$$\bar{C}_p = 2^{d(\frac{1}{2}-\frac{1}{p})}(2\pi)^{\frac{d}{2}}|s(1)|\left(\int_0^\infty \rho^{d-1}\left|\frac{J_{\frac{d}{2}}(\rho)}{\rho^{d/2}}\right|^p\right)^{1/p},$$

where $|s(1)|$ is the surface area of the unit ball in $\mathbb{R}^d$, $d \ge 2$.

Similar estimates may be obtained for the L_p $(1 < p \le \infty)$ norms of the discrete Dirichlet kernel.

Note: These results are particular cases of the so-called Hardy-Littlewood Theorem (see, for instance, Zigmund [76], V. II, XII, §6), which can be formulated as follows:

Theorem B.1. *Let $a_n \ge a_{n+1} \ge \cdots$ and $a_n \to 0$. Consider the series*

$$\sum_{n=1}^{\infty} a_n \cos n\lambda \tag{B.3}$$

and

$$\sum_{n=1}^{\infty} a_n \sin n\lambda \tag{B.4}$$

and define by $f(\lambda)$ and $g(\lambda)$ the sums of the series (B.3) and (B.4) respectively at the points where the series converge. A necessary and sufficient condition that the function f (or g) belongs to L_p, $1 < p < \infty$, is the following

$$\sum_{n=1}^{\infty} a_n^p n^{p-2} \ < \ \infty.$$

Moreover,

$$\|f\|_p^p \asymp \sum_{n=1}^{\infty} a_n^p n^{p-2}.$$

Clearly, $1 = 1 = \cdots = 1 > 0 = \cdots$ is nonincreasing, and we arrive thus to the following estimate for Dirichlet kernels:

$$\| \sum_{t=1}^{T} e^{it\lambda} \|_p \leq C \left(\sum_{t=1}^{T} 1^p t^{p-2} \right)^{\frac{1}{p}} \leq C T^{\frac{p-1}{p}}, 1 < p < \infty,$$

and

$$\int_0^1 \left| \sum_{t=0}^{T-1} e^{2\pi it\lambda} \right|^p d\lambda = T^{p-1} \frac{2}{\pi} \int_0^{\infty} \left| \frac{\sin u}{u} \right|^p du + R_p(T),$$

where the error term

$$R_p(T) = \begin{cases} O_p(T^{p-3}), & p > 3 \\ O(\log T), & p = 3 \\ O_p(1), & 1 < p < 3, \end{cases}$$

where O_p means that constants depend on p.(See, e.g., [1].)

Note that for $p = 1$

$$\int_0^1 \left| \sum_{t=0}^{T-1} e^{2\pi it\lambda} \right| d\lambda \asymp \frac{4}{\pi^2} \log T.$$

Appendix C. The homogeneous Hölder-Young-Brascamp-Lieb inequality

Subtle modifications of the conditions of the Hölder inequality must be made when the arguments of the functions involved are restricted to some subspaces [35]. Starting with Brascamp and Lieb [26] and Lieb [56] (who considered only the case $S = \mathbb{R}$), and following with Ball [17], Barthe [18] and Carlen, Loss and Lieb [29], this generalization of the classical inequalities of Hölder and Young seems to have attained now its definite form in the work of Bennett, Carbery, Christ and Tao [20], [19].

We review now a particular case of this result.

Let

$$x = (x_1, \dots, x_m) \in S^m$$

where S may be either the multidimensional torus, integers or reals

$$S = \begin{cases} [-\pi, \pi]^d \\ \mathbb{Z}^d \\ \mathbb{R}^d \end{cases}$$

endowed with the respective normalized Haar measure $\mu(dx)$.

When $d = 1$, the convergence of integrals of the form:

$$\boxed{\int_{x\in S^m} \frac{\mu(dx_1)...\mu(dx_m)}{l_1(x)^{z_1}...l_k(x)^{z_k}}}$$

where $(l_1, ..., l_k)$ are linear transformations

$$l_j : S^m -> S, \quad l_j(x) =< \alpha_j, x >, \quad j = 1, ..., k$$

and where in the first two cases α_j are supposed to have integer coefficients, is a fundamental question arising in many applications.

Let M denote the matrix with columns $\alpha_j, j = 1, ..., k$. It was for long known to physicists that, when M is fixed, convergence holds for $z = (z_1, ..., z_k)$ belonging to a certain "power counting polytope" PCP (these are relatively similar in all the three cases – see Theorem C.1 below).

It was first noticed in [12] and [13], in the easier case of **unimodal matrices** M, that under the same "power counting conditions" on $z_j = p_j^{-1}$, a Hölder-type inequality

$$(GH) \qquad \left|\int_{S^m} \prod_{j=1}^{k} f_j(l_j(x))\mu(dx)\right| \leq K \prod_{j=1}^{k} \|f_j\|_{p_j}$$

holds, with the powers in (**??**) being replaced by arbitrary functions satisfying integrability conditions $f_j \in L_{p_j}, j = 1, ..., k$, and with $K = 1$. Note that Brascamp and Lieb [26] had already studied the analog harder inequality for general matrices (in the case $S = \mathbb{R}^d$), but without pinpointing exactly the polytope; this was done later by Barthe [18].

For an example, consider the integral

$$J = \int_S \int_S f_1(x_1)f_2(x_2)f_3(x_1 + x_2)f_4(x_1 - x_2)dx_1dx_2$$

where $S = \mathbb{R}$. Here $m = 2,\ k = 4$ and the matrix

$$M = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix}$$

has rank $r(M) = 2$. The theorem below will ensure that

$$|J| \leq \|f_1\|_{1/z_1}\ \|f_2\|_{1/z_2}\ \|f_3\|_{1/z_3}\|f_4\|_{1/z_4}$$

for any $z = (z_1, z_2, z_3, z_4) \in [0, 1]^4$ satisfying $z_1 + z_2 + z_3 + z_4 = 2$, e.g. if $z = (0, 1, 1/2, 1/2)$, then

$$|J| \leq \left(\sup_{0\leq x\leq 1} |f_1(x)|\right) \left(\int_0^1 |f_2(x)|dx\right) \left(\int_0^1 f_3^2(x)dx\right)^{1/2} \left(\int_0^1 f_4^2(x)dx\right)^{1/2}.$$

It is easy to check (and true in general) that the extremal points of the PCP have only 0 and 1 coordinates (which may be exploited for establishing the result). Note also that the matrix M in this example is not unimodal; as a consequence, the optimal constant $K = K(z)$ is not 1 at all the

extremal points, the exception being $(0,0,1,1)$, where it is 2^{-1}; also, the functions achieving equality must be Gaussian (which holds in general, cf. Brascamp-Lieb [26]).

We will formulate now simultaneously the Hölder-Young-Brascamp-Lieb inequality in the three cases:

(C1) $\mu(dx_j)$ is normalized Lebesgue measure on the torus $[-\pi,\pi]^d$, and M has all its coefficients integers.

(C2) $\mu(dx_j)$ is counting measure on $\mathbb{Z}^d$, M has all its coefficients integers, and is unimodular, i.e. all its non-singular minors of dimension $m \times m$ have determinant ± 1.

(C3) $\mu(dx_j)$ is Lebesgue measure on $(-\infty,+\infty)^d$.

The result below specifies the domain of validity of Hölder's inequality, called **power counting polytope**, in terms of linear inequalities involving the rank $r(A)$ of arbitrary subsets A of columns of the matrix M (including the empty set $\emptyset$). It is also possible to express the inequalities in terms of the "dual rank" $r^*(A)$ defined by a dual matrix M^* whose lines are orthogonal to those of M, by using the **duality relation**

$$r^*(A) = |A| - r(M) + r(A^c),\ \forall A$$

Theorem C.1. ***(Homogeneous Hölder-Young-Brascamp-Lieb inequality)***. *Let $l_j(x) = x^t\alpha_j, j = 1,...,k$ be linear functionals $l_j : S^m -> S$ where the space S is either the torus $[-\pi,\pi]^d$, $\mathbb{Z}^d$, or $\mathbb{R}^d$. Let M denote the matrix with columns $\alpha_j, j = 1,...,k$, and let $r(A), r^*(A)$, denote the rank and dual rank of any set A of columns of M.*

Let $f_j,\ j = 1,\dots,k$ be functions $f_j \in L_{p_j}(\mu(dx)),\ 1 \le p_j \le \infty$, defined on S, where $\mu(dx)$ is respectively normalized Lebesgue measure, counting measure and Lebesgue measure.

Let $z_j = \frac{1}{p_j},\ j = 1,\dots,k$, and $z = (z_1,...,z_k)$. The Hölder-Young-Brascamp-Lieb inequality (GH) will hold (with $K = K(z) < \infty$) throughout the "power counting polytopes" PCP defined respectively by:

(c1) $\displaystyle\sum_{j\in A} z_j \le r(A),\ \forall A$

(c2) $\displaystyle\sum_{j\in A} z_j \ge r(M) - r(A^c), \quad \Leftrightarrow \quad \sum_{j\in A}(1-z_j) \le r^*(A)\ \forall A$

(c3) $\displaystyle\sum_{j=1}^{k} z_j = m$, *and one of the conditions (c1) or (c2) is satisfied.*

Alternatively, the conditions (c1-c3) in the theorem are respectively equivalent to:

(1) *$z = (z_1,\dots,z_k)$ lies in the convex hull of the indicators of the sets of independent columns of M, including the void set.*

(2) *$z = (z_1,\dots,z_k)$ lies in the convex hull of the indicators of the sets of columns of M which span its range.*

(3) *$z = (z_1,\dots,z_k)$ lies in the convex hull of the indicators of the sets of columns of M which form a basis.*

If, moreover, the matrix M is unimodal, then the Hölder-Young-Brascamp-Lieb inequality (GH) holds with constant $K = 1$.

Notes: 1) Polytopes defined by the type of rank constraints appearing in cases (c1-c2) are called "polymatroids" (associated to M and M^*) – see Welsh [72], 18.3, Theorem 1. The third polytope is the intersection of the first two.

2) The first two cases of Theorem C.1 were obtained for unimodal matrices in [12] and [13], respectively.

3) Some further important issues, like the precise formula for $K(z)$, and the nonhomogeneous extension where l_j may be linear operators with possibly different images, were resolved only recently – see Lieb [56], Bennett, Carbery, Christ, Tao [20].

Proof sketch: By Edmonds theorem (see Welsh [72], 18.4, Theorem 1) the extremal points of the above polymatroids have only 0 and 1 coordinates for any matrix M. This fact leads to an easy proof, since at the extremal points the result is immediate. For example, in the first and third cases, the extremal points are in one to one correspondence with the indicators of independent sets and bases A, respectively, and the constant at such a point, by a change of variables, is seen to be:

$$K_A = \sqrt{\det(AA^t)}$$

Since $K(z)$ is finite at the extremal points, and Riesz-Thorin interpolation ensures the convexity of $\log(K(z))$, it follows that $K(z)$ will remain finite throughout the polytope generated by the indicators 1_A.

Appendix D. Wick products and Appell polynomials

Let W be a finite set and Y_i, $i \in W$ be a system of random variables. Let

$$Y^W = \prod_{i \in W} Y_i$$

denote the ordinary product, with $Y^\emptyset = 1$, let $m^W = E \prod_{i \in W} Y_i$ be the (mixed) moment, and let

$$\chi(Y^W) = \chi(Y_i, i \in W)$$

denote the (mixed) cumulant of the variables $Y_i, i \in W$, defined recursively as the solutions of the equations:

$$m^W = \sum_{\{V\}|-W} \chi(Y^{V_1}) \cdots \chi(Y^{V_r}), \tag{D.1}$$

where the sum $\sum_{\{V\}|-W}$ is over all partitions $\{V\} = (V_1, \dots, V_r)$, $r \geq 1$ of the set W, and where $\chi(Y^\emptyset) = 1$.

Notes: 1) The equation (D.1) is the formal power series expression of the "exponential relation $m = e^\chi$" between moments and cumulants, viewed as functions on the lattice of subsets [59].

2) The inverse of the equation (D.1), the formal power series expression of the "logarithmic relation $\chi = \log(m)$" may also be computed by:

$$\chi(Y_1, \dots, Y_n) = \frac{\partial^T}{\partial z_1 \dots \partial z_n} \log E \exp(\sum_{i=1}^{T} z_j Y_j)\Big|_{z_1 = \cdots = z_n = 0},$$

where the differentiation is interpreted formally if the moment generating function does not exist.

Definition D.1. The *Wick products* $: Y^W :$ are defined as the solutions of the recursion:

$$Y^W = \sum_{U \subset W} : Y^U : E(Y^{W \setminus U}) = \sum_{U \subset W} : Y^U : \sum_{\{V\} | - W \setminus U} \chi(Y^{V_1}) \cdots \chi(Y^{V_r}),$$

where the sum $\sum_{U \subset W}$ is taken over all subsets $U \subset W$, including $U = \emptyset$, the sum $\sum_{\{V\}|-W\setminus U}$ is over all partitions $\{V\} = (V_1, \ldots, V_r)$, $r \geq 1$ of the set $W \setminus U$, and the starting value is $: Y^{\emptyset} := 1$.

Notes: 1) Inverting the recursion yields ([69], Proposition 1):

$$: Y^W := \sum_{U \subset W} Y^U \sum_{\{V\}|-W\setminus U} (-1)^r \chi(Y^{V_1}) \cdots \chi(Y^{V_r}),$$

as may be formally seen by replacing m^{-1} by $e^{-\chi}$.

2) When some variables appear repeatedly, it is convenient to use the notation

$$: \underbrace{Y_{t_1}, \ldots, Y_{t_1}}_{n_1}, ..., \underbrace{Y_{t_k}, \ldots, Y_{t_k}}_{n_k} := P_{n_1,\ldots,n_k}(Y_{t_1}, ..., Y_{t_k})$$

(the indices in P correspond to the number of times that the variables in ": :" are repeated). The resulting multivariate polynomials $P_{n_1,\ldots,n_k}$ are known as Appell polynomials. These polynomials are a generalization of the Hermite polynomials, which are obtained if Y_t are Gaussian, and like them they play an important role in the limit theory of quadratic forms of dependent variables (cf. [69], [44], [11]).

4) The Appell polynomials may also be directly defined by "power-type" recursions like:

$$\begin{aligned}
&\frac{\partial}{\partial x_j} P_{n_1,\ldots,n_k}(x_1, \ldots, x_k) = n_j P_{n_1,\ldots,n_j - 1,\ldots,n_k}(x_1, \ldots, x_k), \quad E P_{n_1,\ldots,n_k}(X_1, \ldots, X_k) = 0 \\
&\forall n_j \geq 0, j = 1, ..., k, \ \sum_j n_j \geq 1, \\
&P_{0,\ldots,0}(x_1, \ldots, x_k) = 1.
\end{aligned}$$

For example, when $m = n = 1$, $P_{1,1}(X_t, X_s) = X_t X_s - \mathbb{E} X_t\, X_s$, and the bilinear form $Q_T(P_{1,1})$ is a weighted periodogram with its expectation removed.

Note that the multivariate Appell polynomials can be defined by using characteristic functions as well (see, e.g., [70]).

Appendix E. The diagram formula and the moments/cumulants of sums/bilinear forms of Wick products

E.1. **The cumulants diagram representation**

An important property of the Wick products is the existence of simple combinatorial rules for calculation of the (mixed) cumulants, analogous to the familiar diagrammatic formalism for the mixed cumulants of the Hermite polynomials with respect to a Gaussian measure [58]. Let us assume that W is a union of (disjoint) subsets $W_1, \ldots, W_k$. If $(i, 1), (i, 2), \ldots, (i, n_i)$ represent the elements of the

subset W_i, $i = 1, \dots, k$, then we can represent W as a table consisting of rows $W_1, \dots, W_k$, as follows:

$$\begin{pmatrix} (1,1), \dots, (1, n_1) \\ \dots\dots\dots \\ (k,1), \dots, (k, n_k) \end{pmatrix} = W. \tag{E.1}$$

By a *diagram* γ we mean a partition $\gamma = (V_1, \dots, V_r)$, $r = 1, 2, \dots$ of the table W into nonempty sets V_i (the "edges" of the diagram) such that $|V_i| \geq 1$. We shall call the edge V_i of the diagram γ *flat*, if it is contained in one row of the table W; and *free*, if it consists of one element, i.e. $|V_i| = 1$. We shall call the diagram *connected*, if it does not split the rows of the table W into two or more disjoint subsets. We shall call the diagram $\gamma = (V_1, \dots, V_r)$ *Gaussian*, if $|V_1| = \dots = |V_r| = 2$. Suppose given a system of random variables $Y_{i,j}$ indexed by $(i,j) \in W$. Set for $V \subset W$,

$$Y^V = \prod_{(i,j) \in V} Y_{i,j}, \quad \text{and} \quad :Y^V: = :(Y_{i,j}, (i,j) \in V): \ .$$

For each diagram $\gamma = (V_1, \dots, V_r)$ we define the number

$$I_\gamma = \prod_{j=1}^{r} \chi(Y^{V_j}). \tag{E.2}$$

Proposition E.1. *(cf. [44], [69]) Each of the numbers*

$$\begin{array}{ll} (i) & EY^W = E(Y^{W_1} \dots Y^{W_k}), \\ (ii) & E(:Y^{W_1}: \cdots :Y^{W_k}:), \\ (iii) & \chi(Y^{W_1}, \dots, Y^{W_k}), \\ (iv) & \chi(:Y^{W_1}:, \dots, :Y^{W_k}:) \end{array}$$

is equal to

$$\sum I_\gamma,$$

where the sum is taken, respectively, over
(i) all diagrams,
(ii) all diagrams without flat edges,
(iii) all connected diagrams,
(iv) all connected diagrams without flat edges.
If $EY_{i,j} = 0$ for all $(i,j) \in W$, then the diagrams in (i)-(iv) have no singletons.

Notes: 1) Part (i) is just the exponential relation between moments and cumulants.
2) From part (ii) follows, for example, that $E:Y^W: = 0$ (take $W = W_1$, then W has only 1 row and all diagrams have flat edges).

E.2. **Multilinearity**

An important property of Wick products and of cumulants is their multilinearity. For sums and bilinear forms

$$S_T = S_T^m = \int_{I_T} P_m(X_t)\nu(dt), \quad Q_T = Q_T^{m,n} = \int_{I_T}\int_{I_T} \hat{b}(t-s)P_{m,n}(X_t, X_s)\nu(dt)\nu(ds)$$

this implies that:

(1)

$$\chi_k(S_T, ..., S_T) = \int_{t_i \in I_T} \chi(: X_{t_{1,1}}, \ldots, X_{t_{1,m}} :, \ldots, : X_{t_{k,1}}, \ldots, X_{t_{k,m}} :) \prod_{i=1}^{k} \nu(dt_i),$$

where the cumulant in the integral is taken for a table W of k rows $R_1 ...., R_k$, each containing the Wick product of l variables identically equal to X_{t_k}.

(2)

$$\chi_k(Q_T, ..., Q_T) = \int_{t_i, s_i \in I_T} \chi(: X_{t_{1,1}}, \ldots, X_{t_{1,m}},\ X_{s_{1,1}}, \ldots, X_{s_{1,n}} :, \ldots, : X_{t_{k,1}}, \ldots, X_{t_{k,m}},\ X_{s_{k,1}}, \ldots, X_{s_{k,n}} :) \prod_{i=1}^{k} \hat{b}(t_i - s_i) \nu(dt_i) \nu(ds_i),$$

where the cumulant in the integral needs to be taken for a table W of k rows $R_1 ...., R_k$, each containing the Wick product of m variables identically equal to X_{t_k} and of n variables identically equal to X_{s_k}.

E.3. **The cumulants of sums and quadratic forms of moving average tables**

By part (iv) of Proposition E.1, applied to a table W of k rows $R_1 ...., R_k$, with $K = n_1 + ...n_k$ variables, and by the definition (E.2) and of $I\gamma$, we find the following formula for the cumulants of the Wick products of linear variables (4.7):

$$\chi(: X_{t_{1,1}}, \ldots, X_{t_{1,n_1}} :, \ldots, : X_{t_{k,1}}, \ldots, X_{t_{k,n_k}} :) = \sum_{\gamma \in \Gamma(n_1, \ldots, n_k)} \kappa_\gamma J_\gamma(\vec{t}), \tag{E.3}$$

where $\Gamma(n_1, \ldots, n_k)$ denotes the set of all connected diagrams $\gamma = (V_1, \ldots, V_r)$ without flat edges of the table W, $\kappa_\gamma = \chi_{|V_1|}(\xi_{I_1}) \ldots \chi_{|V_r|}(\xi_{I_1})$ and

$$
\begin{aligned}
J_\gamma(t_1, ..., t_K) &= \prod_{j=1}^{r} J_{V_j}(t_{V_j}) \qquad \text{(E.4)}\\
&= \int_{s_1,\ldots,s_r \in\ I} \prod_{j=1}^{k} \Big[\hat{a}(t_{j,1} - s_{j,1})\hat{a}(t_{j,n_1} - s_{j,n_1}) \ldots \\
&\qquad \ldots \hat{a}(t_{k,1} - s_{k,1}) \ldots \hat{a}(t_{k,n_k} - s_{k,n_k})\Big] \nu(ds_1), \ldots, \nu(ds_r)\\
&= \int_{\lambda_1,\ldots,\lambda_K} e^{i \sum_{j=1}^{K} t_j\ \lambda_j} \prod_{i=1}^{K} a(\lambda_i) \prod_{j=1}^{r} \delta(\sum_{i \in V_j} \lambda_i) \prod_{i=1}^{K} \mu(d\lambda_i)\\
&= \int_{\lambda_1,\ldots,\lambda_K} e^{i \sum_{j=1}^{K} t_j\ \lambda_j} \prod_{j=1}^{r} \left(f_{|V_j|}(\lambda_{j,1}, ..., \lambda_{j,|V_j|-1}) \delta(\sum_{i \in V_j} \lambda_i) \right) \prod_{i=1}^{K} \mu(d\lambda_i),
\end{aligned}
$$

where $s_{k,i} \equiv s_j$ if $(k, i) \in V_j$, $j = 1, \ldots, r$ and $\lambda_{j,i} \equiv \lambda_{i+\sum_{l=1}^{j-1} |V_l|}$ if $(j, i) \in V_j$, $j = 1, \ldots, r$.

E.4. **The cumulants of sums and quadratic forms of moving average processes.**

We will apply now the formula (E.4) to compute the cumulants of $S_T^{(m)}, Q_T^{(m,n)}$. In this case, each row j contains just one, respectively two random variables.

It is easy to check that the variance of $S_T^{(2)}$ is:

$$\chi_2(S_T^{(2)}) = 2 \int_{\lambda_1,\lambda_2 \in S} f(\lambda_1) f(\lambda_2) \Delta_T(\lambda_1 - \lambda_2) \Delta_T(\lambda_2 - \lambda_1) \prod_{e=1}^{2} \mu(d\lambda_e).$$

Note that there are only two possible diagrams on a table with two rows of size 2, and that they yield both a graph on two vertices (corresponding to the rows), connected one to the other via two edges.

For another example, the third cumulant $\chi_3(S_T^{(2)})$ is a sum of terms similar to:

$$2^2 \int_{\lambda_1,\lambda_2,\lambda_3 \in S} f(\lambda_1) f(\lambda_2) f(\lambda_3) \Delta_T(\lambda_1 - \lambda_2) \Delta_T(\lambda_2 - \lambda_3) \Delta_T(\lambda_3 - \lambda_1) \prod_{e=1}^{3} \mu(d\lambda_e).$$

This term comes from the 2^2 diagrams in which the row 1 is connected to row 2, 2 to 3 and 3 to 1.

For quadratic forms, a further application of part (iv) of Proposition E.1 will decompose this as a sum of the form

$$\sum_{\gamma \in \Gamma(n_1,\ldots,n_k)} \int_{t_i, s_i \in I_T} R_\gamma(t_i, s_i) \prod_{i=1}^{k} \hat{b}_{t_i - s_i} dt_i ds_i,$$

where $\Gamma(n_1, \ldots, n_k)$ denotes the set of all connected diagrams $\gamma = (V_1, \ldots, V_r)$ without flat edges of the table W and $R_\gamma(t_i, s_i)$ denotes the product of the cumulants corresponding to the partition sets of γ. This easy to check formula is also an illustration of the diagram formula.

When $m = n = 1$ and $k = 2$, besides the Gaussian diagrams we have also one diagram including all the four terms, which makes intervene the fourth order cumulant of X_t.

When $m = n = 1$, the Gaussian diagrams are all products of correlations and the symmetry of $\hat{b}$ implies that all these $2^{k-1}(k-1)!$ terms are equal. We get thus the well-known formula for the cumulants of discrete Gaussian bilinear forms.

In general, we find decompositions as sums of certain "Fejér graph integrals", associated to specific graph structures.

The general structure of the intervening graphs for the cumulants of sums S_T and quadratic forms Q_T have been discussed in Section 3.1 (see Example 3.5).

The following proposition is easy to check.

Proposition E.2. *Let $X_t, t \in I_T$ denote a stationary linear process given by (4.7) with $d = 1$. Then, the cumulants of the sums and quadratic forms defined in (E.4) are given respectively by:*

$$\chi_{k,l} = \chi_k(S_T^{(m)}, ..., S_T^{(m)}) = \sum_{\gamma \in \Gamma(m,k)} \kappa_\gamma \; \sigma_\gamma(T)$$

and

$$\chi_{k,m,n} = \chi_k(Q_T^{(m,n)}, ..., Q_T^{(m,n)}) = \sum_{\gamma \in \Gamma(m,n,k)} \kappa_\gamma \; \tau_\gamma(T),$$

where $\Delta_T(x)$ is the Fejér kernel, $\Gamma(l,k)$, $\Gamma(m,n,k)$ were defined above, and

$$\begin{aligned} \sigma_\gamma(T) &= \int_{\vec{t} \in I_T^k} J_\gamma(\vec{t}) dt \\ &= \int_{\lambda_1,...,\lambda_K} \prod_{j=1}^{k} \Delta_T(\sum_{i=m(j-1)+1}^{m\,j} \lambda_i) \prod_{i=1}^{K} a(\lambda_i) \prod_{j=1}^{r} \delta(\sum_{i \in V_j} \lambda_i) \prod_{i=1}^{K} d\lambda_i, \end{aligned} \tag{E.5}$$

$$\begin{aligned} \tau_\gamma(T) = &\int_{\mu_1,...,\mu_k,\lambda_1,...,\lambda_K,\lambda'_1,...,\lambda'_{K'}} \\ &\prod_{j=1}^{k} \left[\Delta_T\Big(\mu_j + \sum_{i=m(j-1)+1}^{m\,j} \lambda_i\Big) \, \Delta_T\Big(-\mu_j + \sum_{i=n(j-1)+1}^{n\,j} \lambda'_i\Big) \, b(\mu_j) \right] \\ &\times \prod_{i=1}^{K} a(\lambda_i) \prod_{i=1}^{K'} a(\lambda'_i) \prod_{j=1}^{r} \delta(\sum_{i \in V_j} \lambda_i + \sum_{i \in V_j} \lambda'_i) \prod_{i=1}^{K} d\lambda_i \prod_{i=1}^{K'} d\lambda'_i \prod_{i=1}^{k} d\mu_i. \end{aligned} \tag{E.6}$$

These graph structures are simple enough to allow a quick evaluation of the orders of magnitude $\alpha_M(z)$, via the corresponding graph-breaking problems; for the case of bilinear forms we refer to Lemma 1 in [14].

For the case of sums, the domain of applicability of the CLT is $1 - z_1 \geq 1/m$. We check now that at the extremal point $1 - z_1 = 1/m$ we have

$$\begin{aligned}
\alpha_G(z_1) &= \max_A p(A) \\
&= \max_A [co(G-A) - \sum_{e \in A} (1 - z_e)] \\
&= \max_A [co(G-A) - |A|(1 - z_1)] \\
&\leq k/2, \ \forall G \in \mathcal{G}_k,
\end{aligned}$$

where we interpret $p(A)$ as a "profit," equal to the "gain" $co(G-A)$ minus the "cost" $\sum_{e \in A}(1 - z_e)$. We thus need to show that at the extremal point $1 - z_1 = 1/m$,

$$co(G-A)] \leq |A|/m + k/2, \ \forall G \in \mathcal{G}_k.$$

Indeed, this inequality holds with equality for the "total breaking" $A = \mathcal{E}$ (which contains $(km)/2$ edges). It is also clear that no other set of edges A can achieve a bigger "profit" $p(A)$ (defined in (E.7)) than the total breaking, since for any other set A which leaves some vertex still attached to the others, the vertex could be detached from the others with an increase of the number of components by 1 and a cost no more than $m\frac{1}{m}$; thus the profit is nondecreasing with respect to the number of vertices left unattached and thus the total breaking achieves the maximum of $p(A)$.

References

[1] Anderson, B, Ash, J.M., Jones, R., Rider, D.G., and Saffari, B. Exponential sums with coefficients 0 or 1 and concentrated L^p norms. *Annales de l'Institute Fourier,* **57** (2007) 1377-1404.

[2] Anh, V. V. and Leonenko, N.N. Spectral analysis of fractional kinetic equations with random data. *J. Statist. Phys.* **104** (2001) 1349-1387.

[3] Anh, V.V. and Leonenko, N.N. Renormalization and homogenization of fractional diffusion equations with random data. *Prob.Theory and Rel. Fields* **124** (2002) 381-408.

[4] Anh V.V., Angulo, J.M. and Ruiz-Medina, M.D. Possible long-range dependence in fractional random fields, *Journal of Statistical Planning and Inference* **80** (1999) 95–110.

[5] Anh, V.V, Leonenko, N. N., McVinish, R. Models for fractional Riesz-Bessel motion and related processes. *Fractals* **9**, no. 3 (2001) 329–346.

[6] Anh, V.V., Heyde, C.C., Leonenko, N.N. Dynamic models of long-memory processes driven by Lévy noise. *Journal of Applied Probability* **39**, no. 4 (2002) 730–747.

[7] Anh, V.V., Leonenko, N.N., Sakhno, L.M. Quasi-likelihood-based higher-order spectral estimation of random fields with possible long-range dependence. Stochastic Methods and their Applications. *Journal of Applied Probability* **41A** (2004) 35–53.

[8] Anh, V.V., Leonenko, N.N., Sakhno, L.M.: On a class of minimum contrast estimators. *Journal of Statistical Planning and Inference* **123** (2004) 161–185.

[9] Anh, V.V., Leonenko, N. N., Sakhno, L. M. Higher-order spectral densities of fractional random fields. *Journal of Statistical Physics* **111** (2003) 789 – 814.

[10] Avram, F., On Bilinear Forms in Gaussian Random Variables and Toeplitz Matrices. *Probability Theory and Related Fields* **79** (1988) 37–45.

[11] Avram, F., Taqqu, M.S. Noncentral limit theorems and Appell polynomials. *The Annals of Probability* **15** (1987) 767–775.

[12] Avram, F., Brown, L. A Generalized Hölder Inequality and a Generalized Szegö Theorem. *Proceedings of the American Mathematical Society* **107** (1989) 687–695.

[13] Avram, F., Taqqu, M.S. Hölder's Inequality for Functions of Linearly Dependent Arguments. *SIAM Journal of Mathematical Analysis* **20** (1989) 1484–1489.

[14] Avram, F. Generalized Szegö Theorems and asymptotics of cumulants by graphical methods. *Transactions of the American Mathematical Society* **330** (1992) 637–649.

[15] Avram, F., Fox, R. Central limit theorems for sums of Wick products of stationary sequences. *Transactions of the American Mathematical Society* **330** (1992) 651–663.

[16] Avram, F., Taqqu, M.S. On a Szegö type limit theorem and the asymptotic theory of random sums, integrals and quadratic forms. *Dependence in probability and statistics*, Lecture Notes in Statist., 187, Springer, New York, (2006), 259–286.

[17] Ball, K. Volume ratios and a reverse isoperimetric inequality. *J. London Math. Soc. (2)* **44** (1991) 351–359.

[18] Barthe, F. On a reverse form of the Brascamp-Lieb inequality. *Inventiones Mathematicae* **134** (2005) 335–361.

[19] Bennett, J., Carbery, A., Christ, M., Tao, T. Finite bounds for Hölder-Brascamp-Lieb multilinear inequalities. *Math. Res. Lett.,* to appear.

[20] Bennett, J., Carbery, A., Christ, M., Tao, T. The Brascamp-Lieb inequalities: finiteness, structure and extremals, *Geometric and Functional Analysis* (2008), in press.

[21] Bentkus, R.: On the error of the estimate of the spectral function of a stationary process.*Lietuvos Matematikos Rinkinys* **12** no.1 (1972) 55–71 (In Russian).

[22] Bentkus, R. and R. Rutkauskas. On the asymptotics of the first two moments of second order spectral estimators. *Liet. Mat. Rink.* **13(1)** (1973) 29–45.

[23] Beran, J. *Statistics for Long-Memory Processes.* Chapman & Hall, New York (1994).

[24] Bercu, B., Gamboa, F., Lavielle, M. Sharp large deviations for Gaussian quadratic forms with applications. *ESAIM Probabilty and Statistics* **4** (2000) 1–24.

[25] Bercu, B., Gamboa, Rouault, A. Large deviations for quadratic forms of stationary Gaussian processes. *Stochastic Processes and their Applications* **71** (1997) 75–90.

[26] Brascamp, H.J., Lieb E. Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Advances in Math.* **20** (1976) 151–173.

[27] Breuer, P., Major, P. Central limit theorems for nonlinear functionals of gaussian fields. *J. Multiv. Anal.* **13** (1983) 425–441.

[28] Brockwell, P.J. Representations of continuous-time ARMA processes. Stochastic Methods and their Applications. *Journal of Applied Probability* **41A** (2004), 375-382.

[29] Carlen, E.A., Lieb E.H. and Loss, M. A sharp analog of Young's inequality on S^n and related entropy inequalities. *Jour. of Geom. Anal.* **14** (2004) 487–520.

[30] Chambers, D., Slud, E. Central limit theorems for non-linear functions of stationary Gaussian processes. *Probability Theory and Related Fields* **80** (1989) 323–346.

[31] Chambers, D., Slud, E. Necessary conditions for non-linear functions of Gaussian processes to satisfy central limit theorems. *Stochastic Processes and their Applications* **32** (1989) 93-107.

[32] Dobrushin, R. L., Major, P. Non-central limit theorems for non-linear functions of Gaussian fields. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **50** (1979) 27–52.

[33] Fox R., Taqqu, M.S. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics* **14**(2) (1986) 517–532.

[34] Fox, R., Taqqu, M.S. Central limit theorems for quadratic forms in random variables having long-range dependence. *Probability Theory and Related Fields* **74** (1987) 213–240.

[35] E. Friedgut, Hypergraphs, entropy, and inequalities. *Amer. Math. Monthly* **111**, no. 9 (2004) 749-760.

[36] Gay, R. and Heyde, C.C. On a class of random field models which allows long range dependence. *Biometrika* **77** (1990) 401–403.

[37] Gao, J., Anh, V.V., and Heyde, C.C. Statistical estimation of nonstationary Gaussian process with long-range dependence and intermittency, *Stoch. Proc. Appl.* **99** (2002) 295-321.

[38] Ginovian, M.S. On Toeplitz type quadratic functionals of stationary Gaussian processes. *Probability Theory and Related Fields* **100** (1994) 395–406.

[39] Ginovian, M.S. and Sahakyan, A.A. Limit theorems for Toeplitz quadratic functionals of continuous-time stationary processes. *Probability Theory and Related Fields* **138** (2007) 551–579.

[40] Giraitis, L. Central limit theorem for functionals of linear processes. *Lithuanian Math. Journal* **25** (1985) 43-57.

[41] Giraitis, L., and Surgailis, D. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle estimate. *Probab. Theor. Rel. Fields* **86** (1990) 87-104.

[42] Giraitis, L., Taqqu, M. S. Limit theorems for bivariate Appell polynomials, Part 1: Central limit theorems. *Probability Theory and Related Fields* **107** (1997) 359–381.

[43] Giraitis, L., Taqqu, M. S. Whittle estimator for finite variance non-Gaussian time series with long memory. *The Annals of Statistics* **27**(1) (1999) 178–203.

[44] Giraitis, L., Surgailis, D. Multivariate Appell polynomials and the central limit theorem. In E. Eberlein and M. S. Taqqu, editors, *Dependence in Probability and Statistics*, Birkhäuser, New York (1986) 21–71.

[45] Granger, C.W. and Joyeux. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* **10** (1990) 233– 257.

[46] Grenander, V., Szegö, G. *Toeplitz forms and their applications*, Univ. of California Press, Berkeley (1958).

[47] Heyde, C., Gay, R. On asymptotic quasi-likelihood. *Stochastic Processes and their Applications* **31** (1989) 223–236.

[48] Heyde, C., Gay, R. Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. *Stochastic Processes and their Applications* **45** (1993) 169–182.

[49] Heyde, C.C. *Quasi-Likelihood And Its Applications: A General Approach to Optimal Parameter Estimation*. Springer-Verlag, New York (1997).

[50] Hosking, J.R.M. Fractional differencing. *Biometrika* **68** (1981) 165–176.

[51] Hurst, H.E. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* **116** (1951) 770–808.

[52] Ibragimov, I.A. On maximum likelihood estimation of parameters of the spectral density of stationary time series. *Theory Probab. Appl.* **12** (1967) 115-119.

[53] Ivanov, A.V.and Leonenko, N.N. *Statistical Analysis of Random Processes*, Kluwer Academic Publisher, Dordrecht (1989).

[54] Kelbert, M., Leonenko N.N., Ruiz-Medina, M.D. Fractional random fields associated with stochastic fractional heat equation. *Advances of Applied Probability* **37** (2005) 108–133.

[55] Kwapien, S., Woyczynski, W.A. *Random Series and Stochastic Integrals: Single and Multiple*. Birkhaäser, Boston (1992).

[56] E. H. Lieb, Gaussian kernels have only Gaussian maximizers. *Invent. Math.* **102** (1990) 179-208.

[57] Leonenko, N.N. and Sakhno, L.M. On the Whittle estimators for some classes of continuous parameter random processes and fields. *Stat. Probability Letters* **76** (2006) 781-795.

[58] Malyshev, V. A. Cluster expansions in lattice models of statistical physics and the quantum theory of fields. *Russian Mathematical Surveys* **35**(2) (1980) 1–62.

[59] Niven, I. Formal power series. *Amer. Math. Monthly* **76** (1969) 871–889.

[60] D. Nualart and G. Peccati, Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33**, no. 1 (2005), 177–193.

[61] Oxley, J.G. *Matroid Theory*. Oxford University Press, New York (1992).

[62] Oxley, J.G. *What is a matroid*. Preprint, www.math.lsu.edu/ oxley/survey4.pdf (2004).

[63] Peccati G. and C. A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités* XXXVIII (2004) 247-262.

[64] Peligrad, M. and Utev, S. Central limit theorem for stationary linear processes. *Ann. Probab.* **34**, no. 4 (2006) 1608–1622.

[65] Rajput, B. and Rosinski, J. Spectral representations of infinity divisible processes. *Prob. Theory Rel. Fields* **82** (1989) 451-487.

[66] Samorodnitsky, G. and Taqqu, M.S. *Stable Non-Gaussian Random Processes*. Chapman and Hall, New York (1994).

[67] Solev, V., and Gerville-Reache, L. A sufficient condition for asymptotic normality of the normalized quadratic form $\Psi_n(f,g)$. *C.R. Acad. Sci. Paris, Ser. I* **342** (2006) 971-975.

[68] Stein, E.M. *Singular Integrals and Differential Properties of Functions*. Princeton University Press (1970).

[69] Surgailis, D. On Poisson multiple stochastic integral and associated Markov semigroups. Probability and Mathematical Statistics, **3** (1984) 217–239.

[70] Surgailis, D. Long-range dependence and Appel rank, *Ann. Probab.*, **28** (2000) 478-497.

[71] Taqqu, M. S. Convergence of integrated processes of arbitrary Hermite rank, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **50** (1979) 53–83.

[72] Welsh, D. *Matroid Theory*. Academic Press, London (1976).

[73] Willinger, W., Taqqu, M.S., Teverovsky, V. Stock market prices and long-range dependence. *Finance and Stochastics* **3** (1999) 1–13.

[74] Whittle, P. *Hypothesis Testing in Time Series*. Hafner, New York (1951).

[75] Whittle, P. Estimation and information in stationary time series. *Ark. Math.*, **2** (1953) 423-434.

[76] Zygmund, A. *Trigonometric Series*, Vol I, II, third ed., Cambridge Univer. Press (2002).